

























(Lambda ZAP) as a XhoI/SacI insert.  
 gene  
 /ecotype="Columbia"  
 /gene="At5g20000"  
 5'UTR  
 1..78  
 7/gene="At5g20000"  
 /gene="At5g20000"  
 /codel\_start=1  
 /evidence=experimental  
 RPT64  
 /note="putative 26S proteasome AAA-ATPase subunit  
 /protein\_id="AK661442.1"  
 3'UTR  
 1339..1352  
 /gene="At5g20000"  
 /note="3'UTR region in genomic sequence"  
 ORIGIN  
 446 a 278 c 319 g 409 t  
 BASE COUNT  
 Query Match 64.6%; Score 810; DB 8; Length 1512;  
 Best local similarity 80.2%; Pred. No. 6,66-218;  
 Matches 931; Conservative 0; Mismatches 235; Indels 0; Gaps 0;  
 Db 69 GCCGACCAAGAGGCGAAGGCGCTCCGCGACCTACCTATTCCTCAACATGCAAGCAATCA 126  
 150 GAGCGGACAAAGAGGCGAAGGCGCTCCGCGACCTACCTACCTCAAGCTGATGAGAAAGCACT 209  
 Qy 129 GCGCTCTTCGCGAAAGAGGCGCTCCGCGACCTACCTACCTCAAGCTGATGAGAAAGCACT 188  
 Db 210 GCGCTCTTCGCGAAAGAGGCGCTCCGCGACCTACCTACCTCAAGCTGATGAGAAAGCACT 269  
 Qy 189 CAATCTAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTACCTCAAGCAAGCGCTCTTA 248  
 Db 270 CAATCTAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTACCTCAAGCAAGCGCTCTTA 329  
 Qy 249 TCGCGATGATGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 308  
 Db 310 TCGCGATGATGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 369  
 Qy 309 AAGCAAAATGCTGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 368  
 Db 360 AAGCAAAATGCTGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 449  
 Qy 369 TGAAGATGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 428  
 Db 450 GAGATGCTCTTCGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 509  
 Qy 429 TCGATGCTCTTCGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 488  
 Db 510 TCGATGCTCTTCGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 569  
 Qy 489 TCGATGCTCTTCGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 548  
 Db 570 TCGATGCTCTTCGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 629  
 Qy 549 TCGATGCTCTTCGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 608  
 Db 630 TCGATGCTCTTCGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 689  
 Qy 609 TCGATGCTCTTCGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 668  
 Db 690 TCGATGCTCTTCGCAAGGCGAAGGCGAAGGCGCTCCGCGACCTACCTCAAGCAAGCGCTCTTA 749

Qy 669 CAGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 728  
 Db 750 CAGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 809  
 Qy 729 GGAATCTTTTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 786  
 Db 810 GGAATCTTTTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 869  
 Qy 809 GGAATCTTTTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 868  
 Db 870 GGAATCTTTTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 929  
 Qy 849 TACTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 908  
 Db 930 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 989  
 Qy 909 TACTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 968  
 Db 990 TACTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1049  
 Qy 969 AGACGGAATTTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1028  
 Db 1050 TACTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1109  
 Qy 1029 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1088  
 Db 1110 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1169  
 Qy 1089 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1148  
 Db 1170 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1229  
 Qy 1149 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1208  
 Db 1230 TCTGATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1289  
 Qy 1209 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1264  
 Db 1290 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1335  

FIGURE 12  
 AEF220199 1747 bp mRNA linear, plus 26 JAN 2000  
 DEFINITION Pinus taeda 26S proteasome regulatory subunit 8 mRNA, complete cds.  
 ACCESSION AEF220199  
 VERSION AEF220199.1 GI:6752879  
 KEYWORDS  
 SOURCE  
 ORGANISM Pinus taeda (loblolly pine)  
 Pinus taeda, hybrid large; Stereophylla; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus taeda  
 REFERENCES  
 1 (bases 1 to 1747) J.C. Casae, Mollano A. and Cairney J. Pinus taeda, a full-length cDNA encoding a putative 26S proteasome subunit 8 from Pinus taeda somatic embryos  
 2 (bases 1 to 1747)  
 JOURNAL Deletio Belian, J.C., Casae, Mollano A. and Cairney J.  
 AUTHORS Submitted (30-DEC-1999) Forest Biology, Institute of Paper Science and Technology, 500 10th St. NW, Atlanta, GA 30318, USA  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /accession="AF220199"  
 /gi="6752879"  
 /clone="pRBS11.1"  
 /tissue="somatic embryo"  
 18..1335  
 /note="similar to Tac-binding protein; AAA-ATPase family"















[illegible]





PR 16-APR-1999 99US-0130077  
 PR 21-APR-1999 99US-0130449  
 PR 23-APR-1999 99US-0130510  
 PR 23-APR-1999 99US-0130891  
 PR 28-APR-1999 99US-0131449  
 PR 28-APR-1999 99US-0131450  
 PR 30-APR-1999 99US-0132407  
 PR 04-MAY-1999 99US-0132484  
 PR 05-MAY-1999 99US-0132485  
 PR 06-MAY-1999 99US-0132487  
 PR 07-MAY-1999 99US-0132631  
 PR 11-MAY-1999 99US-0132636  
 PR 11-MAY-1999 99US-0132618  
 PR 14-MAY-1999 99US-0132719  
 PR 14-MAY-1999 99US-0132721  
 PR 18-MAY-1999 99US-0133769  
 PR 18-MAY-1999 99US-0133769  
 PR 20-MAY-1999 99US-0135114  
 PR 21-MAY-1999 99US-0135353  
 PR 24-MAY-1999 99US-0135639  
 PR 27-MAY-1999 99US-0136021  
 PR 28-MAY-1999 99US-0136782  
 PR 01-JUN-1999 99US-0137222  
 PR 02-JUN-1999 99US-0137258  
 PR 07-JUN-1999 99US-0137254  
 PR 08-JUN-1999 99US-0138094  
 PR 10-JUN-1999 99US-0138540  
 PR 14-JUN-1999 99US-0139119  
 PR 16-JUN-1999 99US-0139452  
 PR 16-JUN-1999 99US-0139453  
 PR 16-JUN-1999 99US-0139454  
 PR 18-JUN-1999 99US-0139455  
 PR 18-JUN-1999 99US-0139456  
 PR 18-JUN-1999 99US-0139457  
 PR 18-JUN-1999 99US-0139459  
 PR 18-JUN-1999 99US-0139460  
 PR 18-JUN-1999 99US-0139461  
 PR 18-JUN-1999 99US-0139463  
 PR 18-JUN-1999 99US-0139760  
 PR 18-JUN-1999 99US-0139763  
 PR 23-JUN-1999 99US-0140353  
 PR 23-JUN-1999 99US-0140354  
 PR 24-JUN-1999 99US-0140635  
 PR 24-JUN-1999 99US-0140635  
 PR 29-JUN-1999 99US-0140691  
 PR 30-JUN-1999 99US-0141287  
 PR 01-JUL-1999 99US-0141842  
 PR 02-JUL-1999 99US-0142065  
 PR 06-JUL-1999 99US-0142300  
 PR 08-JUL-1999 99US-0142803  
 PR 12-JUL-1999 99US-0143200  
 PR 13-JUL-1999 99US-0143452  
 PR 14-JUL-1999 99US-0143524  
 PR 15-JUL-1999 99US-0144005  
 PR 16-JUL-1999 99US-0144086  
 PR 19-JUL-1999 99US-0144325  
 PR 19-JUL-1999 99US-0144331  
 PR 19-JUL-1999 99US-0144333  
 PR 19-JUL-1999 99US-0144334  
 PR 19-JUL-1999 99US-0144334

PR 19-JUL-1999 99US-0144335  
 PR 20-JUL-1999 99US-0144332  
 PR 20-JUL-1999 99US-0144684  
 PR 20-JUL-1999 99US-0144684  
 PR 21-JUL-1999 99US-0145088  
 PR 21-JUL-1999 99US-0145088  
 PR 22-JUL-1999 99US-0145087  
 PR 22-JUL-1999 99US-0145089  
 PR 22-JUL-1999 99US-0145132  
 PR 23-JUL-1999 99US-0145219  
 PR 23-JUL-1999 99US-0145224  
 PR 26-JUL-1999 99US-0145276  
 PR 27-JUL-1999 99US-0145913  
 PR 27-JUL-1999 99US-0145919  
 PR 28-JUL-1999 99US-0145931  
 PR 28-AUG-1999 99US-0145866  
 PR 02-AUG-1999 99US-0146089  
 PR 03-AUG-1999 99US-0147038  
 PR 04-AUG-1999 99US-0147044  
 PR 05-AUG-1999 99US-0147052  
 PR 05-AUG-1999 99US-0147600  
 PR 06-AUG-1999 99US-0147603  
 PR 09-AUG-1999 99US-0147116  
 PR 09-AUG-1999 99US-0147335  
 PR 10-AUG-1999 99US-0148171  
 PR 11-AUG-1999 99US-0148319  
 PR 13-AUG-1999 99US-0148654  
 PR 13-AUG-1999 99US-0148654  
 PR 15-AUG-1999 99US-0149368  
 PR 16-AUG-1999 99US-0149370  
 PR 16-AUG-1999 99US-0149420  
 PR 20-AUG-1999 99US-0149723  
 PR 20-AUG-1999 99US-0149723  
 PR 23-AUG-1999 99US-0149920  
 PR 23-AUG-1999 99US-0149920  
 PR 25-AUG-1999 99US-0150565  
 PR 26-AUG-1999 99US-0150684  
 PR 27-AUG-1999 99US-0151066  
 PR 27-AUG-1999 99US-0151066  
 PR 30-AUG-1999 99US-0151103  
 PR 01-SEP-1999 99US-0151130  
 PR 01-SEP-1999 99US-0152633  
 PR 07-SEP-1999 99US-0152633  
 PR 10-SEP-1999 99US-0153070  
 PR 12-SEP-1999 99US-0153769  
 PR 12-SEP-1999 99US-0153769  
 PR 16-SEP-1999 99US-0154423  
 PR 20-SEP-1999 99US-0154779  
 PR 22-SEP-1999 99US-0155133  
 PR 22-SEP-1999 99US-0155133  
 PR 24-SEP-1999 99US-0155459  
 PR 24-SEP-1999 99US-0156589  
 PR 29-SEP-1999 99US-0156596  
 PR 05-OCT-1999 99US-0157117  
 PR 06-OCT-1999 99US-0157665  
 PR 07-OCT-1999 99US-0158029  
 PR 12-OCT-1999 99US-0158232  
 PR 13-OCT-1999 99US-0159293  
 PR 13-OCT-1999 99US-0159294  
 PR 14-OCT-1999 99US-0159295  
 PR 14-OCT-1999 99US-0159325





PR 23-APR-1993 9905-01300921  
 PR 28-APR-1993 9905-0131449  
 PR 30-APR-1993 9905-0120208  
 PR 01-MAY-1993 9905-0132000  
 PR 04-MAY-1993 9905-0132844  
 PR 05-MAY-1993 9905-0132485  
 PR 06-MAY-1993 9905-0132486  
 PR 07-MAY-1993 9905-0132481  
 PR 11-MAY-1993 9905-0134256  
 PR 14-MAY-1993 9905-0134219  
 PR 14-MAY-1993 9905-0134218  
 PR 14-MAY-1993 9905-0134370  
 PR 18-MAY-1993 9905-0134768  
 PR 20-MAY-1993 9905-0134941  
 PR 21-MAY-1993 9905-0135353  
 PR 24-MAY-1993 9905-0135649  
 PR 27-MAY-1993 9905-0136021  
 PR 28-MAY-1993 9905-0136782  
 PR 28-MAY-1993 9905-0137522  
 PR 01-JUN-1993 9905-0137222  
 PR 03-JUN-1993 9905-0137228  
 PR 07-JUN-1993 9905-0137724  
 PR 08-JUN-1993 9905-0138094  
 PR 10-JUN-1993 9905-0138480  
 PR 10-JUN-1993 9905-0139452  
 PR 14-JUN-1993 9905-0139119  
 PR 15-JUN-1993 9905-0139452  
 PR 15-JUN-1993 9905-0139453  
 PR 16-JUN-1993 9905-0139454  
 PR 18-JUN-1993 9905-0139455  
 PR 18-JUN-1993 9905-0139456  
 PR 18-JUN-1993 9905-0139456  
 PR 18-JUN-1993 9905-0139459  
 PR 18-JUN-1993 9905-0139460  
 PR 18-JUN-1993 9905-0139462  
 PR 18-JUN-1993 9905-0139462  
 PR 21-JUN-1993 9905-0139817  
 PR 22-JUN-1993 9905-0139859  
 PR 23-JUN-1993 9905-0140351  
 PR 24-JUN-1993 9905-0140351  
 PR 24-JUN-1993 9905-0140351  
 PR 28-JUN-1993 9905-0140823  
 PR 29-JUN-1993 9905-0140991  
 PR 01-JUL-1993 9905-0141481  
 PR 01-JUL-1993 9905-0142154  
 PR 02-JUL-1993 9905-0142055  
 PR 06-JUL-1993 9905-0142803  
 PR 06-JUL-1993 9905-0142803  
 PR 09-JUL-1993 9905-0142820  
 PR 12-JUL-1993 9905-0142877  
 PR 14-JUL-1993 9905-0143354  
 PR 15-JUL-1993 9905-0144065  
 PR 15-JUL-1993 9905-0144085  
 PR 16-JUL-1993 9905-0144086  
 PR 16-JUL-1993 9905-0144331  
 PR 19-JUL-1993 9905-0144332  
 PR 19-JUL-1993 9905-0144333  
 PR 19-JUL-1993 9905-0144335  
 PR 19-JUL-1993 9905-0144335  
 PR 20-JUL-1993 9905-0144632

PR 20-JUL-1993 9905-0144884  
 PR 21-JUL-1993 9905-0144814  
 PR 21-JUL-1993 9905-0145086  
 PR 22-JUL-1993 9905-0145086  
 PR 22-JUL-1993 9905-0145087  
 PR 22-JUL-1993 9905-0145089  
 PR 23-JUL-1993 9905-0145174  
 PR 23-JUL-1993 9905-0145218  
 PR 23-JUL-1993 9905-0145218  
 PR 26-JUL-1993 9905-0145276  
 PR 26-JUL-1993 9905-0145271  
 PR 27-JUL-1993 9905-0145519  
 PR 27-JUL-1993 9905-0145519  
 PR 28-JUL-1993 9905-0145525  
 PR 28-JUL-1993 9905-0145525  
 PR 02-AUG-1993 9905-0146388  
 PR 03-AUG-1993 9905-0147028  
 PR 03-AUG-1993 9905-0146389  
 PR 04-AUG-1993 9905-0147502  
 PR 05-AUG-1993 9905-0147192  
 PR 05-AUG-1993 9905-0147560  
 PR 06-AUG-1993 9905-0147516  
 PR 06-AUG-1993 9905-0147516  
 PR 09-AUG-1993 9905-0147935  
 PR 09-AUG-1993 9905-0147935  
 PR 11-AUG-1993 9905-0148311  
 PR 11-AUG-1993 9905-0148311  
 PR 12-AUG-1993 9905-0149785  
 PR 12-AUG-1993 9905-0149785  
 PR 13-AUG-1993 9905-0149785  
 PR 13-AUG-1993 9905-0149785  
 PR 16-AUG-1993 9905-0149786  
 PR 17-AUG-1993 9905-0149785  
 PR 18-AUG-1993 9905-0149785  
 PR 20-AUG-1993 9905-0149785  
 PR 20-AUG-1993 9905-0149785  
 PR 23-AUG-1993 9905-0149785  
 PR 23-AUG-1993 9905-0149785  
 PR 25-AUG-1993 9905-0149785  
 PR 26-AUG-1993 9905-0149785  
 PR 27-AUG-1993 9905-0149785  
 PR 27-AUG-1993 9905-0149785  
 PR 30-AUG-1993 9905-0151305  
 PR 31-AUG-1993 9905-0151305  
 PR 01-SEP-1993 9905-0151305  
 PR 01-SEP-1993 9905-0151305  
 PR 10-SEP-1993 9905-0153070  
 PR 13-SEP-1993 9905-0153755  
 PR 15-SEP-1993 9905-0154218  
 PR 15-SEP-1993 9905-0154218  
 PR 20-SEP-1993 9905-0147787  
 PR 22-SEP-1993 9905-0155133  
 PR 23-SEP-1993 9905-0155133  
 PR 26-SEP-1993 9905-0156566  
 PR 29-SEP-1993 9905-0156596  
 PR 04-OCT-1993 9905-0171117  
 PR 04-OCT-1993 9905-0171162  
 PR 07-OCT-1993 9905-0176022  
 PR 09-OCT-1993 9905-0186172  
 PR 13-OCT-1993 9905-0192992  
 PR 13-OCT-1993 9905-0192992  
 PR 13-OCT-1993 9905-0192992  
 PR 14-OCT-1993 9905-0193920  
 PR 14-OCT-1993 9905-0193920  
 PR 14-OCT-1993 9905-0193920



















253 GGTGAGTGTGTAAGATGGATGGGACAGAGCAAGTCCCTGTCAAGTCCACCGAGAGG 312  
 Db 1999 GCGAGATGGTGAAGTCCAGTCCAGTCCAGAGAGAGTGTGTGAGATGATCTATCCGAGGCG 1346  
 Cy 313 AAATATGTGTGATGATGAGTCAAGAAATATGATGATCTCAAGATCTCCACCTCA 372  
 Db 1939 AAGATATGTATGTATGTATGAGTCAAGATCAAGTCAATTTAAAGATATGACACAGAGACG 1880  
 Cy 373 GTTCCACCTCCGACCAAGCAAGATGTATGTTCTCTCACTGTGTGTGCGACATGAATTTTCA 432  
 Db 1879 GTGCGCCCTCCGACCAAGCAAGTGTATGTTCTCTCACTGTGTGTGCGACATGAATTTTCA 1820  
 Cy 433 TTGTGTATATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 492  
 Db 1819 TTATATCCCTTATGTGTGTGTGAGAAAGTGTGCGCAATCTCCACCTCAAGAAATGTGTGTG 1760  
 Cy 493 TTATGACGCGAATTTAAAGATATGATGATGATGATGATGATGATGATGATGATGATG 552  
 Db 1759 TTGTGACAGCAGATGATGAGTATACAGAGAGATGATGATGATGATGATGATGATGATG 1700  
 Cy 553 GTTGTATGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 612  
 Db 1699 TTATATGACGCGAATTTAAAGATATGATGATGATGATGATGATGATGATGATGATG 1640  
 Cy 613 AAGATGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 672  
 Db 1639 AAGATGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1580  
 Cy 672 GTTGT 732  
 Db 1579 GTTGT 1520  
 Cy 732 GTTGT 792  
 Db 1519 GTTGT 1460  
 Cy 792 GTTGT 852  
 Db 793 ATTCGATGTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 852  
 Db 1459 ATTCGATGTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1406  
 Cy 852 ATTCGATGTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 912  
 Db 1406 ATTCGATGTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1346  
 Cy 912 ATTCGATGTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 972  
 Db 1346 ATTCGATGTCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1286  
 Cy 972 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1012  
 Db 1286 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1226  
 Cy 1012 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1092  
 Db 1226 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1166  
 Cy 1092 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1152  
 Db 1166 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1106  
 Cy 1152 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1212  
 Db 1106 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1046  
 Cy 1212 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1233  
 Db 1046 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1005

RESULT 14  
 AB232038  
 ID AB232038 standard; NM; 1206 BP.

AC AB232038;  
 XX 30-JAN-2003 (first entry)  
 DT  
 DB Candida albicans essential gene seq ID NO 6325.  
 XX  
 XX Fungus; yeast; tetracycline; promoter; GRAC acetal; biosynthesis;  
 XX signal transduction; DNA replication; cell division; growth  
 XX proliferation; candida albicans; fungus; cell; antifungal; gene; ss.  
 XX  
 OS Candida albicans.  
 XX  
 XX W02001053728-42.  
 XX  
 XX 11-JUL-2002.  
 PD  
 XX 26-DEC-2001; 2001NC-USA9586.  
 XX  
 XX 29-DEC-2000; 2000US-2591287.  
 PR  
 XX 20-FEB-2001; 2001US-0792024.  
 PR  
 XX 22-AUG-2001; 2001US-314050F.  
 XX  
 PA (EUTR-) EUTRA PHARM INC.  
 XX  
 XX Roemer T, Jiang B, Boone C, Bussey H, Olsen KJ,  
 XX WPI: 2002-566694/60.  
 DR  
 XX P-RSDB; ABP73488.  
 XX  
 XX Constructing strains for identifying gene products as effective targets  
 PT for therapeutic intervention, by inactivating in the seven one allele  
 PT of a gene and placing other allele of the gene under conditional  
 XX expression  
 XX  
 XX Claim 37; SEQ ID NO 6325; 167bp + Sequence Listing; English.  
 XX  
 XX The invention relates to constructing (W1) a strain of diploid fungus  
 CC one allele by deletion or replacement by a cassette having an  
 CC repressible selectable marker and modifying other allele by  
 CC recombination, of a promoter replacement fragment with a heterologous  
 CC promoter, (W2) a method for constructing a strain of diploid fungus  
 CC cells in which both alleles of a gene are modified. The diploid fungus  
 CC cells having both alleles modified are useful for identifying a gene that  
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene  
 CC that contributes to the resistance of a diploid fungus to an antifungal  
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus  
 CC disease, (W3) is useful for identifying a compound which modulates the  
 CC activity of a gene product, preferably enzymatic activity, carbon  
 CC compound catabolism, biosynthesis, transporter, transcriptional,  
 CC translational, signal transduction, DNA replication and cell division  
 CC regulation, cell growth, cell division, cell death, cell survival, cell  
 CC ability to inhibit growth or proliferation of C. albicans cells and for  
 CC treating infection by C. albicans. The present sequence is that of an  
 CC essential candida albicans gene used in the method of the invention.  
 CC The sequence is a partial sequence of a gene product, identified by  
 CC specification but is based on sequence information supplied to Devent by  
 CC the European Patent Office.  
 CC  
 XX Sequence 1206 BP; 425 A; 170 C; 259 G; 352 T; 0 other;  
 XX  
 XX Query Match 42.8%; Score 536.6; DB 24; Length 1206;  
 XX Best Local Similarity 66.6%; Pval No. 5.1e-156;  
 XX Matches 793; Complement 389; Indels 3; Gaps 1;  
 XX  
 XX 79 GAGGCGAGGAGCGCTCCGACATCTCTCAATCCAGAGCATCAATCACTCTCTCT 138  
 Db 31 CTGAGAGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 90  
 Cy 139 GCGCAAGAGCTCAATCACTCTCCAGGCTCCAGAGGAAACAGATCAATTCAG 198

[illegible]

**RESULTS**

AerBt1919 standard; DNW, 1170 BP.

AerBt1919:

Aspergillus fumigatus essential gene H1557.

Fungalicide; cytoskeletal; essential gene; Aspergillus fumigatus; infection;

cancer; contamination; biofilm; antibody; immune response; ds.

Aspergillus fumigatus.

Molecular cloning and sequencing analysis of the Aspergillus fumigatus  
essential genes of Aspergillus fumigatus.

The invention relates to novel purified or isolated nucleic acids of  
essential genes of Aspergillus fumigatus. The isolated nucleic acid of  
organism such as A. fumigatus, or treat non-infectious disease in a subject  
subject (e.g., cancer), to prevent or contain contamination of an object  
by A. fumigatus, or to prevent or inhibit formation on a surface of a  
expressing recombinant protein from host cells.  
therapeutic use, as markers for those tissues in which the pathogenic  
organisms invade or reside, for comparing with the DNA sequence of A.  
strain or bioprocessed product, for identifying polynucleotide sequences  
potential orthologous essential or virulence genes, for selecting and  
of drug candidates for attachment to a nucleic acid array for examination  
anti-gen to raise anti-DNA antibodies or to elicit another immune re-  
response, and for identifying polymorphonuclear encoding the other proteins  
with which binding occurs or to identify inhibitors of the binding  
elicited immune response, as a reagent in assays designed to quantitat-

determine levels of the protein in biological fluids, as a marker for  
host tissues in which pathogenic organisms invade or reside, and to  
factors. This polynucleotide sequence represents one of the essential  
genes of Aspergillus fumigatus of the invention.

Sequence 1170 BP : 311 A: 267 C: 322 G: 170 T: 0 other:  
Query Match      40.7%, Score 510.4: DB 25, Length 1170:

Best Local Similarity 65.4%; Pred. No. 7,2e-148;  
Matches 764; Conservative 0; Mismatches 401; Indels 3; Gaps 1

[illegible][illegible]

Search completed: November 9, 2003, 09:03:39  
Job time : 446 secs

C	28	143.11	3376	US-09-149-476-298	Sequence 298, App
C	29	140.98	11	US-08-812-1790-1	Sequence 1, Appl
C	31	135.2	10.8	US-09-216-338-106	Sequence 106, App
C	32	135.2	10.8	US-09-216-338-106	Sequence 106, App
C	33	128	10.2	US-09-157-468-11	Sequence 1, Appl
C	34	128	10.2	US-09-157-468-11	Sequence 1, Appl
C	35	127.8	10.2	US-09-816-093-1	Sequence 1, Appl
C	36	122.4	9.8	US-09-328-017-794	Sequence 794, App
C	37	118.2	9.8	US-09-328-017-794	Sequence 794, App
C	38	118.2	9.8	US-09-328-017-794	Sequence 794, App
C	39	112.4	9.0	US-09-103-840A-1	Sequence 1, Appl
C	40	112.4	9.0	US-09-103-840A-1	Sequence 1, Appl
C	41	98.4	7.8	US-09-313-294K-1390	Sequence 1390, Ap
C	42	99.4	7.8	US-08-933-028-1176	Sequence 1176, Ap
C	43	96.2	7.7	US-09-313-294K-1379	Sequence 4379, Ap
C	44	96.2	7.7	US-09-313-294K-1379	Sequence 4379, Ap
C	45	95.4	7.6	US-08-998-416-372	Sequence 372, App

  

		ALIGNMENTS	
<hr/>			
RESULTS			
US-08-916-421B-1			
Sequence 1, Application US/08916421B			
Patent No. 6503729			
APPLICANT INFORMATION: 1.			
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methano-			
Patent No. 6503729			
FILE OF INVENTION: jamaasch1			
CURRENT FILING DATE: 1997-08-22			
PRIOR APPLICATION NUMBER: US/08/916,428			
NUMBER OF SEQ ID NOS: 3			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO 1			
TYPE DNA			
SEQUENCE 1644976			
ORGANISM: Methanococcus jamaaschii			
FEATURES: misc-feature			
LOCATION: (12822)..(18221)			
OTHER INFORMATION: n equals a, t, c, or g			
NAME/KEY: misc-feature			
LOCATION: (12825)..(18218)			
OTHER INFORMATION: n equals a, t, c, or g			
NAME/KEY: misc-feature			
LOCATION: (84773)..(84773)			
OTHER INFORMATION: n equals a, t, c, or g			
NAME/KEY: misc-feature			
LOCATION: (16460)..(16460)			
OTHER INFORMATION: n equals a, t, c, or g			
NAME/KEY: misc-feature			
LOCATION: (16481)..(16481)			
OTHER INFORMATION: n equals a, t, c, or g			
NAME/KEY: misc-feature			
LOCATION: (98120)..(98120)			
OTHER INFORMATION: n equals a, t, c, or g			
NAME/KEY: misc-feature			
LOCATION: (98159)..(98159)			
OTHER INFORMATION: n equals a, t, c, or g			
NAME/KEY: misc-feature			
LOCATION: (98261)..(98261)			
OTHER INFORMATION: n equals a, t, c, or g			
NAME/KEY: misc-feature			
LOCATION: (98313)..(98313)			

















[illegible]

```

1  RESULT 10
2  US-09-273-565-15
3  Application US/09273565A
4  Patent No. 7,166,190
5  GENERAL INFORMATION:
6  APPLICANT: FUJIMURA, TSUTOMU
7  INVENTOR: FUJIMURA, TSUTOMU
8  APPLICANT: KOSHI, SHUNO
9  TITLE OF INVENTION: AN ISOLATED MUSCLE-SPECIFIC
10  MUSCULIN ENCODING HUMAN
11  FILE REFERENCE: 0-53199
12  CURRENT FILING DATE: 1999-03-22
13  EARLIER APPLICATION NUMBER: 09/055,199
14  EARLIER FILING DATE: 1998-04-07/820,170
15  EARLIER FILING DATE: 1997-03-19
16  EARLIER APPLICATION NUMBER: JP 63410/1996
17  EARLIER FILING DATE: 1996-03-15/66163/1997
18  EARLIER FILING DATE: 1997-03-06
19  NUMBER OF SEQ ID NOS: 95
20  SOFTWARE: PatentIn Ver. 2.1
21  SEQ ID NO 1566
22  TYPE: DNA
23  ORGANISM: Homo sapiens
24  FEATURE:
25  NAME: CDS
26  LOCATION: (17)..(1183)
27  US-09-273-565-15

```

[illegible]

RESULT 11  
US-09-565-538-15  
; Sequence 15, Application US/09565538  
; Patent No. 6333404  
; GENERAL INFORMATION:  
; APPLICANT: FUJIMURA, TSUTOMU  
; APPLICANT: WATANABE, TAKESHI



APPLICANT: HORIE, MASATO  
 TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN  
 SKELETAL MUSCLE-SPECIFIC UR-QUITIN-CONJUGATING ENZYME  
 FILE REFERENCE: O-51599  
 CURRENT APPLICATION NUMBER: US/09/565,538  
 CURRENT FILING DATE: 2000-03-05  
 PRIOR FILING DATE: 1996-03-13  
 PRIOR FILING DATE: 1999-03-22  
 PRIOR APPLICATION NUMBER: 09/055,699  
 PRIOR FILING DATE: 1996-04-07  
 PRIOR APPLICATION NUMBER: 08/820,170  
 PRIOR FILING DATE: 1996-03-13  
 PRIOR APPLICATION NUMBER: JP 63110/1996  
 PRIOR FILING DATE: 1996-03-13  
 PRIOR APPLICATION NUMBER: JP 69463/1997  
 PRIOR FILING DATE: 1997-03-05  
 NUMBER OF SEQ ID NOS: 5  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 15  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1566  
 ORIGIN: 1566  
 US-09-565-538-15  
 Query Match 21.6% Score 270.6, DB 4, Length 1566;  
 Best Local Similarity 55.3%; Pied No. 2,96-744  
 Matches 535; Conservative 0; Mismatches 419; Indels 3; Gaps 1;

250 GTCGTGAGAAAGTTGGTGAAGAATGATGAGGAAGAAAGTCTTGGTCAAGTCCAGCCGAA 309  
 185 GGTGGTGGTGGTCTTAACTGACCTACTAGAGAAATCTGTTAACTGACATGGA 344  
 310 GGAATAATGTTGTGATGACCAAAATATTTGATCCAAAGATATCTTCATCACT 429  
 245 CAGAGATGTTGTGATGTTGTTCTGTCACAGCTGACAGAAAGTACAGCAGAGACA 304  
 370 AAGATGATGATCCGACAGACAGATGATATGTTCTGACTATGTTCCAGTAAAGTTAT 429  
 305 AAGATGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 364  
 430 CCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489  
 365 CCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424  
 490 GATTGACACCAATTTAAAGAAATTAAGAGATGATGATGATGATGATGATGATGAT 549  
 425 GAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 484  
 550 GAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609  
 485 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544  
 610 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669  
 545 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 604  
 670 AAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729  
 605 AAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664  
 730 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789  
 665 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 724  
 790 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849  
 725 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781  
 850 ACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909

Db 782 ACGTATGAGAGTTACGTATGATCAATGATGATGATGATGATGATGATGATGATGAT 841  
 910 TTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 969  
 Db 842 ACGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901  
 970 GACGAGAAATTTATTTTCAACCCCTATATGATGATGATGATGATGATGATGAT 1029  
 Db 902 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 961  
 1010 CATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069  
 Db 962 CTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021  
 1090 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1149  
 Db 1022 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081  
 1150 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206  
 Db 1082 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138  
 RESULT 12  
 Sequence 15, Application us/09661468  
 Patent No. 6376189  
 GENERAL INFORMATION:  
 APPLICANT: HORIE, MASATO  
 APPLICANT: HORIE, MASATO  
 APPLICANT: HORIE, MASATO  
 TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN  
 SKELETAL MUSCLE-SPECIFIC UR-QUITIN-CONJUGATING ENZYME  
 FILE REFERENCE: O-51599  
 CURRENT APPLICATION NUMBER: US/09/565,538  
 CURRENT FILING DATE: 2000-03-13  
 PRIOR FILING DATE: 1996-03-13  
 PRIOR FILING DATE: 1999-03-22  
 PRIOR APPLICATION NUMBER: 09/055,699  
 PRIOR FILING DATE: 1996-04-07  
 PRIOR APPLICATION NUMBER: 08/820,170  
 PRIOR FILING DATE: 1996-03-13  
 PRIOR APPLICATION NUMBER: JP 63110/1996  
 PRIOR FILING DATE: 1996-03-13  
 PRIOR APPLICATION NUMBER: JP 69463/1997  
 PRIOR FILING DATE: 1997-03-05  
 NUMBER OF SEQ ID NOS: 5  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 1566  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1566  
 ORIGIN: 1566  
 US-09-561-468-15  
 Query Match 21.6% Score 270.6, DB 4, Length 1566;  
 Best Local Similarity 55.3%; Pied No. 2,96-744  
 Matches 535; Conservative 0; Mismatches 419; Indels 3; Gaps 1;

250 GTCGTGAGAAAGTTGGTGAAGAATGATGAGGAAGAAAGTCTTGGTCAAGTCCAGCCGAA 309  
 185 GGTGGTGGTGGTCTTAACTGACCTACTAGAGAAATCTGTTAACTGACATGGA 344  
 310 GGAATAATGTTGTGATGACCAAAATATTTGATCCAAAGATATCTTCATCACT 429  
 245 CAGAGATGTTGTGATGTTGTTCTGTCACAGCTGACAGAAAGTACAGCAGAGACA 304  
 370 AAGATGATGATCCGACAGACAGATGATATGTTCTGACTATGTTCCAGTAAAGTTAT 429  
 305 AAGATGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 364  
 430 CCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489  
 365 CCAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 424  
 490 GATTGACACCAATTTAAAGAAATTAAGAGATGATGATGATGATGATGATGATGAT 549  
 425 GAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 484  
 550 GAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609  
 485 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544  
 610 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669  
 545 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 604  
 670 AAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729  
 605 AAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664  
 730 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789  
 665 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 724  
 790 AAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849  
 725 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 781  
 850 ACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 899





Db 181 AAGTCATTCAGAAAAATGAACTTGAATGCGCGCATTCGAAAAAGATCGCGGAAA 240  
 Or 1085 AATGCAATCGACATCGCTCGAACTTAAAGCTT 1121  
 Db 241 AATGCAATCGCGCTTCGAACTGAACTCGAAAGGCGCT 277  
 Search completed: November 9, 2003, 11:16:59  
 codb time : 116 secs



QY	136	CTTCGCGAAGAGCTCATATCTCAACGCTTCGCGAGGTGTAGAGAAACACCTCAATCT	195
DB	268	GGATATATATAGAGCGAAACAGATATATCACTCTTCCGAGACCCGAGTACCAATTAATCT	327
QY	196	AAGGTGATGATATAGAGCGAAACAGATATATCACTCTTCCGAGACCCGAGTACCAATTAATCT	255
DB	338	AAATGTGACCTATTTCGGAGAGAGGTACACCTGCTCCGAGACCGGGCTCTATGTGCGG	387
QY	256	GAATGTTCATGATTAAGTGTGACGAGACCAATCTCTGTCAAGCTCAACCGAGAGAA	315
DB	388	GAATGTTCATGATTAAGTGTGACGAGACCAATCTCTGTCAAGCTCAACCGAGAGAA	447
QY	316	TATCTGTGTATCACTAATGCAAAATTTATATCTCAAGATGTCTCCATCAACATGATTA	375
DB	448	TTGTGTGTGATGATGAGCAAAAACCTTATCACTACAGATGTGTACCAAGTCCCGAGT	503
QY	376	GCATCTCGCAAGACAGATATGTCTTCTCATGATCTCTCCGAGTAAATGATCAACAT	435
DB	508	GCTCTATGATTAATGATACACCTGACACCTGACATCTCCGAGACAGATGATGATTA	567
QY	436	GTCTATGATTAATGATACACCTGACACCTGACATCTCCGAGACAGATGATGATTA	495
DB	568	GTCTATGATTAATGATACACCTGACACCTGACATCTCCGAGACAGATGATGATTA	627
QY	496	GACGACGATATTTAAATTAATTAAGAGGTGTATGGCTACACATCAACGTCGAGT	555
DB	628	GATACGATTAATGATTAATTAAGAGGTGTATGGCTACACATCAACGTCGAGT	687
QY	556	TTTGTGATCTTATGATATCTCAACAGATAGATGATGATGATCTCTGATGAGACCTG	615
DB	688	TTTGTGATCTTATGATATCTCAACAGATAGATGATGATGATCTCTGATGAGACCTG	747
QY	616	GCTTAAATATGT	675
DB	748	GCTTAAATATGT	807
QY	676	TCTGT	735
DB	808	TCTGT	867
QY	736	TTTGT	795
DB	868	TTTGT	927
QY	796	GT	854
DB	928	GT	985
QY	886	CTGTGATCTTCAACGCTGATATGATATGTGATGACTTCAATAGATCAAGGTTTTAT	915
DB	986	CTGTGATCTTCAACGCTGATATGATATGTGATGACTTCAATAGATCAAGGTTTTAT	1044
QY	916	GTGTGATCTTCAACGCTGATATGATATGTGATGACTTCAATAGATCAAGGTTTTAT	975
DB	1045	GTGTGATCTTCAACGCTGATATGATATGTGATGACTTCAATAGATCAAGGTTTTAT	1104
QY	976	AAAAATTTTCTCAACCTCTTATATAGAGTCTGTGTGTGTGTGTGTGTGTGTGTGT	1035
DB	1105	AAAAATTTTCTCAACCTCTTATATAGAGTCTGTGTGTGTGTGTGTGTGTGTGTGT	1164
QY	1036	AAATATATTTTCTCAACCTCTTATATAGAGTCTGTGTGTGTGTGTGTGTGTGTGT	1095
DB	1165	AAATATATTTTCTCAACCTCTTATATAGAGTCTGTGTGTGTGTGTGTGTGTGTGT	1224
QY	1096	GCATCTGT	1155
DB	1225	GCATCTGT	1284
QY	1156	GTGTGATCTTCAACGCTGATATGATATGTGATGACTTCAATAGATCAAGGTTTTAT	1215
DB	1285	GTGTGATCTTCAACGCTGATATGATATGTGATGACTTCAATAGATCAAGGTTTTAT	1344

[illegible]















909 TTGATGCGCCAAATCGATGATGATGCTGGATCGAAGCCCTCTTAAAGACGAGACGAT 968  
 DB 834 GATCCTGCTGACACGACACGACACGACACGACACGACACGACACGACACGACACGAC 993  
 969 AAGCGGAAATGATGATTCGACCCCTTAAAGAGGCTCGGATGATGATGATGATGAT 1028  
 DB 894 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 953  
 1029 CATTCTAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088  
 DB 954 CATTCTAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1013  
 1089 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148  
 DB 1014 AAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073  
 1149 GAGGAGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206  
 DB 1074 TCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131  
  
 RESULT 12  
 US-10-032-585-6660      Application US/1003585  
 Publication No. US20030180953A1  
 GENERAL INFORMATION:  
 APPLICANT: Teiry, Ronger D.  
 APPLICANT: Charles Boone  
 APPLICANT: Howard, Bussey  
 TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
 FILE REFERENCE: US-10-032-5995/10/032,585  
 CURRENT FILING DATE: 2001-12-20  
 NUMBER OF SEQ ID NOS: 8000  
 SOFTWARE: Patent In version 3.1  
 SEQ ID NO: 100  
 TYPE: DNA  
 LENGTH: 1398  
 ORGANISM: Candida albicans  
 US-10-032-585-6660  
  
 Query Match      21.9%; Score 275.2; DB 12; Length 1398;  
 Best Local Similarity 58.8%; Pval No. 1.3e-71;  
 Matches 494; Conservative 0; Mismatches 343; Indels 3; Gaps 1;  
 DB 409 GTTCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 466  
 DB 487 GTTCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546  
 469 TCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528  
 DB 547 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588  
 529 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 566  
 607 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666  
 589 GTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 646  
 667 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726  
 DB 607 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666  
 649 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708  
 727 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786  
 709 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 768  
 787 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846  
 DB 787 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846  
 847 GTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 903

829 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888  
 DB 904 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 963  
 889 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 948  
 DB 949 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1023  
 964 GTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008  
 DB 1024 TTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1083  
 1009 CCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068  
 DB 1084 AAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1143  
 1069 AAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128  
 DB 1144 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203  
 1129 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1188  
 DB 1204 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263  
 1189 ATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1248  
 DB 1244 TCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1123  
  
 RESULT 13  
 US-10-101-510-379      Application US/10101510  
 Sequence 379; Publication No. US20030180953A1  
 GENERAL INFORMATION:  
 APPLICANT: MAN, JACSON  
 APPLICANT: MAN, YIXIN  
 TITLE OF INVENTION: GENE DISRUPTION PROFILES AND METHODS OF USE  
 FILE REFERENCE: 1517,002  
 CURRENT FILING DATE: 2002-03-20  
 NUMBER OF SEQ ID NOS: 6576,947  
 PRIOR FILING DATE: 2001-03-20  
 NUMBER OF SEQ ID NOS: 805  
 SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO: 100  
 TYPE: DNA  
 LENGTH: 1891  
 ORGANISM: Homo sapiens  
 US-10-101-510-379  
  
 Query Match      21.7%; Score 272.2; DB 12; Length 1891;  
 Best Local Similarity 56.0%; Pval No. 1.2e-70;  
 Matches 536; Conservative 0; Mismatches 418; Indels 3; Gaps 1;  
 DB 250 GTTCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 309  
 DB 308 GTTCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 367  
 310 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387  
 DB 338 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447  
 370 AAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429  
 DB 448 AAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507  
 430 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489  
 DB 508 CCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567  
 480 GTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 549





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

ON nucleic - nucleic search, using sw model

Run on: November 9, 2003, 08:53:50 / Search time 2025 seconds

10075.310 Million cell updates/sec

File: us-03-462-972-1

Sequence: 1 atgcctcctcagagagctga.....cattcagagagctctgagag 1254

Scoring table: IDENTITY, NUC

Gapop 10.0, Gapext 1.0

Searched: 2270192 seqs, 1212238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0

Database: Maximum Match 1000

Listing first 45 summaries

1: em. eatba.\*  
2: em. eatba.\*  
3: em. eatba.\*  
4: em. eatba.\*  
5: em. eatba.\*  
6: em. eatba.\*  
7: em. eatba.\*  
8: em. eatba.\*  
9: em. eatba.\*  
10: em. eatba.\*  
11: em. eatba.\*  
12: em. eatba.\*  
13: em. eatba.\*  
14: em. eatba.\*  
15: em. eatba.\*  
16: em. eatba.\*  
17: em. eatba.\*  
18: em. eatba.\*  
19: em. eatba.\*  
20: em. eatba.\*  
21: em. eatba.\*  
22: em. eatba.\*  
23: em. eatba.\*  
24: em. eatba.\*  
25: em. eatba.\*  
26: em. eatba.\*  
27: em. eatba.\*  
28: em. eatba.\*  
29: em. eatba.\*  
30: em. eatba.\*

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	DB ID	Description
1	66.06	54.2	1655 11	AY103724
2	65.32	54.2	1655 11	AY103724
3	63.2	50.4	723 9	AY103724
4	630.2	50.3	1302 11	AK010505

5	632.6	49.6	1649 11	CA784321
6	614	49.0	742 14	CB84406
7	608.4	48.5	794 12	B1366213
8	608.4	48.2	794 12	B1366213
9	605.48	48.2	685 10	BB943284
10	604	48.1	113 11	AK013294
11	602.8	48.1	715 12	AK013294
12	597.6	46.9	629 12	B1364104
13	587.6	46.9	629 12	B1364104
14	587.6	46.9	629 12	B1364104
15	580.8	46.3	632 9	AM85884
16	580.8	46.3	632 9	AM85884
17	579.8	46.1	632 9	AM85884
18	569.4	46.4	976 10	BE145520
19	563.6	44.9	976 10	BE145520
20	562	44.8	1162 14	CD305627
21	562	44.8	1162 14	CD305627
22	560	44.7	560 12	BM731600
23	557	44.4	569 12	B1364538
24	555	44.2	589 12	B1364538
25	555	44.2	589 12	B1364538
26	550	43.9	582 13	BM553255
27	547	43.6	547 12	BM573432
28	546.4	43.6	572 12	B1674291
29	541.8	43.2	541 12	B1674291
30	541.8	43.2	541 12	B1674291
31	541.4	43.2	543 12	B1364533
32	541.4	43.2	543 12	B1364533
33	539.6	42.9	539 12	BM553255
34	539.6	42.9	539 12	BM553255
35	537.4	42.9	603 9	AM939662
36	533	42.5	638 12	B1418891
37	532	42.4	1166 14	CD305627
38	528.8	42.2	733 12	B1364533
39	528.8	42.2	733 12	B1364533
40	526.4	42.0	697 10	BE145871
41	525	41.9	907 10	BE145871
42	525	41.9	907 10	BE145871
43	523	41.7	755 10	BE530658
44	520	41.5	559 12	BE530658
45	516.8	41.2	896 10	BE146483

## ALIGNMENTS

RESULT 1  
LOCUS: AY103724  
DEFINITION: Zea mays PCD078180 mRNA sequence.  
VERSION: AY103724.1 GI:1106802  
KEYWORDS: mRNA, Zea mays, PCD078180  
SOURCE: Zea mays  
ORGANISM: Zea mays  
HTC: 1655 bp mRNA linear HTC 16-OCT-2002

REFERENCE  
JOURNAL: Zea mays PCD078180 mRNA sequence.  
TITLE: Zea mays PCD078180 mRNA sequence.  
AUTHORS: Haney, C.F., Dolan, M., Mao, G.H., Vogel, J.M., Malster, M.S., Arthur, L.M., Haney, M., Morgan, M., and Tingey, S.V.  
Zeal Mapping Project/Dupont Consensus Sequences for Design of Unpublished (2002)

REFERENCE  
JOURNAL: Zea mays PCD078180 mRNA sequence.  
TITLE: Zea mays PCD078180 mRNA sequence.  
AUTHORS: Haney, C.F., Dolan, M., Mao, G.H., Vogel, J.M., Malster, M.S., Arthur, L.M., Haney, M., Morgan, M., and Tingey, S.V.  
Zeal Mapping Project/Dupont Consensus Sequences for Design of Unpublished (2002)

COMMENT: Zea mays PCD078180 mRNA sequence.  
Zeal Mapping Project/Dupont Consensus Sequences for Design of Unpublished (2002)

www.1gtr.org or NCBI, www.ncbi.nlm.nih.gov. When the source of the































```

Db      488 NATTAAGAAATTTTTTGGATGTCCTTGAAGTTACCAATCCAAACATCTGACCTGTTTACAG 547
Oy      564 TCTTGAAATTGCACACCAAGAGTGTCTGCTTATGAGGCTACCTGATACGATMAAC 633
Db      548 TCTTGAAATTGCACATCAAAATGATGTCTCTCTATGAGGCTACCTGATACGATMAAC 607
Oy      624 ATTGATGCTAGAGGAGATGAGCTTAT 648
Db      608 ATTGATGCTAGAGGAGATGAGCTTAT 632

Search completed: November 9, 2003, 11:15:00
Job time : 3031 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2003, 11:15:06 Search time 81 seconds

W/BLAST alignment 819.108 Million cell updates/sec

US-09-462-972-2

Sequence: 1 VALVPELHMAAEVPEAN.....VAKWKEKEMSLKMK 418

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Search: 157263 seqs, 15726373 residues

Total number of hits satisfying chosen parameters: 110763

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database:

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT\*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT\*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT\*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT\*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT\*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT\*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT\*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT\*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT\*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT\*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT\*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT\*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT\*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT\*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT\*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT\*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT\*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT\*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT\*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT\*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT\*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT\*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT\*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT\*
- 25: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2004.DAT\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the local score distribution.

## SUMMARIES

Result	Score	Query	Length	DB	ID	Description
1	2116	100.0	418	20	AAW7650	soybean SUG1 polyp
2	1906.5	90.1	419	21	AAQ4452	Arabidopsis thaliana
3	1896.5	89.1	420	21	AAQ4453	Arabidopsis thaliana
4	1891	89.4	405	21	AAQ4454	Arabidopsis thaliana
5	1860	87.9	392	20	AAW7651	Corn SUG1 polypept
6	1740	82.2	350	21	AAQ4455	Arabidopsis thaliana
7	1730.5	81.7	350	21	AAQ4456	Arabidopsis thaliana
8	1670.5	78.9	402	23	AAH1340	Human coarctin anti
9	1655	78.7	406	20	AAW7653	Trn-interacting UJA

10	1665	78.7	406	20	AAW83374	Human UJA protein
11	1665	78.7	406	20	AAW83389	Human UJA protein
12	1601.5	78.5	405	22	AAH58177	Drosophila melanog
13	1596.5	78.2	405	22	AAH58178	Drosophila melanog
14	1654	78.2	406	15	AAH53158	Thyroid hormone re
15	1622.5	76.7	399	22	AAH58875	Drosophila melanog
16	1540.5	72.8	401	23	AAH73488	Candida albicans e
17	1539.5	72.8	389	24	AAH25111	Aspergillus fumig
18	1509.5	71.2	389	24	AAH25111	Aspergillus fumig
19	1501	70.9	321	21	AAH28315	Arabidopsis thaliana
20	1497.5	70.8	405	15	AAH53529	S. cerevisiae SUG1
21	1497.5	70.8	405	15	AAH53529	S. cerevisiae SUG1
22	1497.5	70.8	405	15	AAH53529	S. cerevisiae SUG1
23	1394	65.9	299	21	AAH28316	Arabidopsis thaliana
24	1105	52.2	241	21	AAH28317	Arabidopsis thaliana
25	97.5	44.3	399	22	AAH56764	Mouse P. ayei
26	97.5	44.3	399	22	AAH56764	Mouse P. ayei
27	90.1	42.6	440	23	AAH5208	Mouse ischaemic co
28	89.4	42.3	428	21	AAH52876	Drosophila melanog
29	89.4	42.3	428	21	AAH52876	Drosophila melanog
30	89.4	42.3	428	21	AAH52876	Drosophila melanog
31	88.5	41.9	346	21	AAH52877	Arabidopsis thaliana
32	88.4	41.8	433	22	AAH58036	Drosophila melanog
33	88.3	41.7	347	21	AAH56968	Arabidopsis thaliana
34	88.3	41.7	347	21	AAH56968	Arabidopsis thaliana
35	87.7	41.4	369	18	AAH51505	Human 26S proteas
36	87.4	41.3	451	24	AAH26313	Aspergillus fumig
37	86.9	41.1	397	22	AAH59155	Drosophila melanog
38	86.8	41.0	393	24	AAH25763	Aspergillus fumig
39	86.8	41.0	393	24	AAH25763	Aspergillus fumig
40	86.8	41.0	399	21	AAH53066	Arabidopsis thaliana
41	86.2	40.7	501	21	AAH52881	Arabidopsis thaliana
42	86.2	40.7	501	21	AAH52881	Arabidopsis thaliana
43	86.2	40.7	501	21	AAH52881	Arabidopsis thaliana
44	86.0	40.6	399	21	AAH52859	Arabidopsis thaliana
45	86.0	40.6	439	24	AAH25864	Aspergillus fumig

## ALIGNMENTS

RESULT 1  
AAW7650 standard; Protein: 418 AA.

ID	AAW7650	soybean SUG1 polypeptide.
1	AAW7650	soybean SUG1 polypeptide.
2	AAW7650	soybean SUG1 polypeptide.
3	AAW7650	soybean SUG1 polypeptide.
4	AAW7650	soybean SUG1 polypeptide.
5	AAW7650	soybean SUG1 polypeptide.
6	AAW7650	soybean SUG1 polypeptide.
7	AAW7650	soybean SUG1 polypeptide.
8	AAW7650	soybean SUG1 polypeptide.
9	AAW7650	soybean SUG1 polypeptide.
10	AAW7650	soybean SUG1 polypeptide.
11	AAW7650	soybean SUG1 polypeptide.
12	AAW7650	soybean SUG1 polypeptide.
13	AAW7650	soybean SUG1 polypeptide.
14	AAW7650	soybean SUG1 polypeptide.
15	AAW7650	soybean SUG1 polypeptide.
16	AAW7650	soybean SUG1 polypeptide.
17	AAW7650	soybean SUG1 polypeptide.
18	AAW7650	soybean SUG1 polypeptide.
19	AAW7650	soybean SUG1 polypeptide.
20	AAW7650	soybean SUG1 polypeptide.
21	AAW7650	soybean SUG1 polypeptide.
22	AAW7650	soybean SUG1 polypeptide.
23	AAW7650	soybean SUG1 polypeptide.
24	AAW7650	soybean SUG1 polypeptide.
25	AAW7650	soybean SUG1 polypeptide.
26	AAW7650	soybean SUG1 polypeptide.
27	AAW7650	soybean SUG1 polypeptide.
28	AAW7650	soybean SUG1 polypeptide.
29	AAW7650	soybean SUG1 polypeptide.
30	AAW7650	soybean SUG1 polypeptide.
31	AAW7650	soybean SUG1 polypeptide.
32	AAW7650	soybean SUG1 polypeptide.
33	AAW7650	soybean SUG1 polypeptide.
34	AAW7650	soybean SUG1 polypeptide.
35	AAW7650	soybean SUG1 polypeptide.
36	AAW7650	soybean SUG1 polypeptide.
37	AAW7650	soybean SUG1 polypeptide.
38	AAW7650	soybean SUG1 polypeptide.
39	AAW7650	soybean SUG1 polypeptide.
40	AAW7650	soybean SUG1 polypeptide.
41	AAW7650	soybean SUG1 polypeptide.
42	AAW7650	soybean SUG1 polypeptide.
43	AAW7650	soybean SUG1 polypeptide.
44	AAW7650	soybean SUG1 polypeptide.
45	AAW7650	soybean SUG1 polypeptide.

Human UJA protein  
Thyroid hormone re  
Drosophila melanog  
Drosophila melanog  
Thyroid hormone re  
Drosophila melanog  
Candida albicans e  
Aspergillus fumig  
Aspergillus fumig  
Arabidopsis thaliana  
SUG1 transcription  
S. cerevisiae SUG1  
S. cerevisiae SUG1  
Arabidopsis thaliana  
Arabidopsis thaliana  
Mouse P. ayei  
Mouse ischaemic co  
Drosophila melanog  
Arabidopsis thaliana  
Arabidopsis thaliana  
Drosophila melanog  
Arabidopsis thaliana  
Human 26S proteas  
Aspergillus fumig  
Drosophila melanog  
Aspergillus fumig  
Arabidopsis thaliana  
Arabidopsis thaliana  
Arabidopsis thaliana  
Arabidopsis thaliana  
Arabidopsis thaliana  
Aspergillus fumig





























GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

On protein - protein search, using sw model

Run on: November 9, 2003, 12:44:17, Search time 28 Seconds

Without alignments 631.641 Million cell updates/sec

US-09-462-972-2

Sequence: 1 MALVQELKHAAGGPGAN.....VAKKKEKTERMSIRKWK 118

Scoring table: BLOSUM62

gapop 10.0, Gapex 0.5

Searched: 328717 seqs, 42310856 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :  
1: Issued Patent, AA\*  
2: /cgn2\_6/prodata/1/aa/58\_COMB\_Pep\*  
3: /cgn2\_6/prodata/1/aa/58\_COMB\_Pep\*  
4: /cgn2\_6/prodata/1/aa/58\_COMB\_Pep\*  
5: /cgn2\_6/prodata/1/aa/58\_COMB\_Pep\*  
6: /cgn2\_6/prodata/1/aa/58\_COMB\_Pep\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	1665	78.7	406 2 US-08-222-719-1	Sequence 1, Appl
2	1665	78.7	406 2 US-08-470-925-1	Sequence 1, Appl
3	1665	78.7	406 2 US-08-470-925-1	Sequence 1, Appl
4	1665	78.7	406 2 US-08-470-925-1	Sequence 1, Appl
5	1497.5	70.8	405 2 US-08-222-719-2	Sequence 2, Appl
6	1497.5	70.8	405 2 US-08-470-925-2	Sequence 2, Appl
7	1497.5	70.8	405 2 US-08-470-925-2	Sequence 2, Appl
8	1497.5	70.8	405 2 US-08-470-925-2	Sequence 2, Appl
9	940.5	44.4	215 4 US-09-291-170A-6	Sequence 6, Appl
10	940.5	44.4	215 4 US-09-291-170A-6	Sequence 6, Appl
11	940.5	44.4	215 4 US-09-291-170A-6	Sequence 6, Appl
12	940.5	44.4	215 4 US-09-291-170A-6	Sequence 6, Appl
13	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
14	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
15	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
16	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
17	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
18	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
19	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
20	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
21	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
22	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
23	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
24	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
25	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
26	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl
27	877	41.4	388 3 US-08-924-884-6	Sequence 13, Appl

28	467.5	22.1	516 4 US-09-724-884-1	Sequence 1, Appl
29	465	22.0	214 4 US-09-291-170A-7	Sequence 7, Appl
30	465	22.0	214 4 US-09-724-884-1	Sequence 7, Appl
31	465	22.0	214 4 US-09-724-884-1	Sequence 7, Appl
32	465	22.0	214 4 US-09-724-884-1	Sequence 7, Appl
33	447	21.5	308 4 US-09-119-476-607	Sequence 607, Appl
34	447	21.1	408 4 US-09-107-5326-6233	Sequence 12, Appl
35	433	20.5	686 2 US-08-993-228-12	Sequence 11, Appl
36	433	20.5	686 2 US-08-993-228-12	Sequence 11, Appl
37	407.5	19.6	220 4 US-09-290-179-4	Sequence 4, Appl
38	407.5	19.6	220 4 US-09-290-179-4	Sequence 4, Appl
39	407.5	19.6	220 4 US-09-290-179-4	Sequence 4, Appl
40	388.5	18.7	361 4 US-09-724-884-4	Sequence 4, Appl
41	388.5	18.7	361 4 US-09-724-884-4	Sequence 4, Appl
42	388	18.3	227 4 US-09-291-170A-9	Sequence 9, Appl
43	388	18.3	227 4 US-09-291-170A-9	Sequence 9, Appl
44	388	18.3	227 4 US-09-291-170A-9	Sequence 9, Appl
45	388	18.3	227 4 US-09-291-170A-9	Sequence 9, Appl
46	388	18.3	227 4 US-09-291-170A-9	Sequence 9, Appl
47	388	18.3	227 4 US-09-291-170A-9	Sequence 9, Appl
48	388	18.3	227 4 US-09-291-170A-9	Sequence 9, Appl
49	388	18.3	227 4 US-09-291-170A-9	Sequence 9, Appl
50	388	18.3	227 4 US-09-291-170A-9	Sequence 9, Appl

ALIGNMENTS

RESULT: 1  
US-08-222-719-1  
Sequence 1, Application US/08222719  
Patent No. 546711  
GENERAL INFORMATION:  
APPLICANT: JAVAD MOORE  
INVENTOR: JAVAD MOORE  
TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR-INTERACTING POLYPEPTIDES AND  
NUMBER OF SEQUENCES: RELATED MOLECULES AND METHODS  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIF: 02110-2804  
COMPUTER: IBM PS/2 Model 502 of 555K  
OPERATING SYSTEM: MS-DOS (version 5.0)  
CURRENT APPLICATION DATA: (Version 5.1)  
APPLICATION NUMBER: US/08222719  
FILING DATE: 04 April 1994  
PRIORITY DATE: 07/969,136  
APPLICATION NUMBER: 1992  
ATTORNEY/INFORMATION:  
NAME: Paul T. Clark  
REGISTRATION NUMBER: 30,162  
RESERVATION/DOCKET NUMBER: 00766/229001  
TELEPHONE: (617) 543-5070  
TELEFAX: (617) 543-8996  
INVENTOR: JAVAD MOORE  
SEQUENCE CHARACTERISTICS:  
LENGTH: 406  
TYPER: amino acid  
TOPOLOGY: linear  
US-08-222-719-1

Query Match 78.7% Score 1665, DB 2  
Best Local Similarity 78.5% Pred. No. 5,967-145  
Matches 38; Conservative 38; Mismatches 40; Indels 12; Gaps 2;

1 MALVUVELKMAHEVPEVNCASFKPTQCGZARVYSLSNHEHQULLRQKNTNANLPAOR 60  
 1 MALDOPEGMELREG-----KASGSLQAOVYLSAIEQLVINDSGSLRRLAQOR 49  
 61 NDLSNRYMAREEQLQLOPGSVGVGVVWQGNVLVWVHEKGVVVDIDVYDITXIT 120  
 50 NBLKNAVYALREELQLOPGSVGVGVVWQGNVLVWVHEKGVVVDIDVYDITXIT 109  
 121 NRYVMAKNDVYALREELQLOPGSVGVGVVWQGNVLVWVHEKGVVVDIDVYDITXIT 180  
 110 PNCVMAKNDVYALREELQLOPGSVGVGVVWQGNVLVWVHEKGVVVDIDVYDITXIT 169  
 181 KPELPEALQAGKQVGLVPGPGTQGLTLAAVAHHTDCTTRVSGSELVQVFGESK 240  
 170 KPELPEALQAGKQVGLVPGPGTQGLTLAAVAHHTDCTTRVSGSELVQVFGESK 229  
 241 WRELFVMAERHASTIPMOEISIGASRNESGNGDSQGVCTMELNLOJCFEASIK 300  
 230 WRELFVMAERHASTIPMOEISIGASRNESGNGDSQGVCTMELNLOJCFEASIK 288  
 301 KVIAMATNRIDILQALRPFIDKLEPTPEPNEBSRDLTKHSRNMVANGDILAKIA 360  
 289 KVIAMATNRIDILQALRPFIDKLEPTPEPNEBSRDLTKHSRNMVANGDILAKIA 348  
 361 EKNVAGSGLAVACTEAGFAPRVRVYTOEDFEMAVAKYNKTEKNSJRLKLR 418  
 349 EKNVAGSGLAVACTEAGFAPRVRVYTOEDFEMAVAKYNKTEKNSJRLKLR 406

RESULT 2  
 US-08-470-925-1  
 / Sequence 1, Application US/08470925  
 / Patent No. 5362256  
 / General Information:  
 / APPLICANT: David D. Moore  
 / TITLE OF INVENTION: NICKELAR HORMONE RECEPTOR-  
 / TITLE OF INVENTION: INTERSECTING POLYPEPTIDES AND  
 / NUMBER OF INVENTION: RELATED MOLECULES AND METHODS  
 / NUMBER OF SEQUENCES: 31  
 / CORRESPONDENCE ADDRESS: 31  
 / STREET: 225 Franklin Street  
 / CITY: Boston  
 / STATE: Massachusetts  
 / COUNTRY: U.S.A.  
 / ZIP: 02110-2804  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: 3.5 Diskette, 1.44 MB  
 / OPERATING SYSTEM: MS-DOS (Version 5.0)  
 / SOFTWARE: WordPerfect (Version 5.1)  
 / CURRENT APPLICATION DATA: 08/470,925  
 / FILING DATE: 06-June-1995  
 / CLASSIFICATION: 435  
 / PRIOR APPLICATION DATA: 08/222,719  
 / FILING DATE: 04-Apr-11-1994  
 / CLASSIFICATION: 435  
 / APPLICATION NUMBER: 07/969,136  
 / OPERATING SYSTEM: MS-DOS (Version 5.0)  
 / SOFTWARE: WordPerfect (Version 5.1)  
 / CURRENT APPLICATION DATA: 08/470,925  
 / FILING DATE: 06-June-1995  
 / CLASSIFICATION: 435  
 / NAME: Paul T. Clark 30, 152  
 / ADDRESS: 225 Franklin Street  
 / TELEPHONE: (617) 542-5070  
 / TELECOMMUNICATION INFORMATION:  
 / TELEFAX: 200134 542-8906  
 / INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
 / Length: 406  
 / TYPE: amino acid  
 / STRANDEDNESS:  
 / ORIENTATION: N/A  
 / US-08-470-925-1  
 Query Match 78.3% Score 1665; DB 2; Length 406;  
 Blast Local Similarity 78.3% Pct Id 5.9e-148; Indels 12; Gaps 2;  
 Matches 328; Mismatch 40;  
 1 MALVUVELKMAHEVPEVNCASFKPTQCGZARVYSLSNHEHQULLRQKNTNANLPAOR 60  
 1 MALDOPEGMELREG-----KASGSLQAOVYLSAIEQLVINDSGSLRRLAQOR 49  
 61 NDLSNRYMAREEQLQLOPGSVGVGVVWQGNVLVWVHEKGVVVDIDVYDITXIT 120  
 50 NBLKNAVYALREELQLOPGSVGVGVVWQGNVLVWVHEKGVVVDIDVYDITXIT 109  
 121 NRYVMAKNDVYALREELQLOPGSVGVGVVWQGNVLVWVHEKGVVVDIDVYDITXIT 180  
 110 PNCVMAKNDVYALREELQLOPGSVGVGVVWQGNVLVWVHEKGVVVDIDVYDITXIT 169  
 181 KPELPEALQAGKQVGLVPGPGTQGLTLAAVAHHTDCTTRVSGSELVQVFGESK 240  
 170 KPELPEALQAGKQVGLVPGPGTQGLTLAAVAHHTDCTTRVSGSELVQVFGESK 229  
 241 WRELFVMAERHASTIPMOEISIGASRNESGNGDSQGVCTMELNLOJCFEASIK 300  
 230 WRELFVMAERHASTIPMOEISIGASRNESGNGDSQGVCTMELNLOJCFEASIK 288  
 301 KVIAMATNRIDILQALRPFIDKLEPTPEPNEBSRDLTKHSRNMVANGDILAKIA 360  
 289 KVIAMATNRIDILQALRPFIDKLEPTPEPNEBSRDLTKHSRNMVANGDILAKIA 348  
 361 EKNVAGSGLAVACTEAGFAPRVRVYTOEDFEMAVAKYNKTEKNSJRLKLR 418  
 349 EKNVAGSGLAVACTEAGFAPRVRVYTOEDFEMAVAKYNKTEKNSJRLKLR 406

RESULT 3  
 US-08-471-613-1  
 / Sequence 1, Application US/08471613  
 / Patent No. 5362256  
 / General Information:  
 / APPLICANT: David D. Moore  
 / TITLE OF INVENTION: NICKELAR HORMONE RECEPTOR-  
 / TITLE OF INVENTION: INTERSECTING POLYPEPTIDES AND  
 / NUMBER OF INVENTION: RELATED MOLECULES AND METHODS  
 / NUMBER OF SEQUENCES: 31  
 / CORRESPONDENCE ADDRESS: 31  
 / STREET: 225 Franklin Street  
 / CITY: Boston  
 / STATE: Massachusetts  
 / COUNTRY: U.S.A.  
 / ZIP: 02110-2804  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: 3.5 Diskette, 1.44 MB  
 / OPERATING SYSTEM: IBM PS/2 Model 50C or 555X  
 / SOFTWARE: WordPerfect (Version 5.1)  
 / CURRENT APPLICATION DATA: 08/471,613  
 / FILING DATE: 06-June-1995  
 / CLASSIFICATION: 435  
 / PRIOR APPLICATION DATA: 08/222,719  
 / FILING DATE: 04-Apr-11-1994  
 / CLASSIFICATION: 435  
 / APPLICATION NUMBER: 07/969,136  
 / OPERATING SYSTEM: MS-DOS (Version 5.0)  
 / SOFTWARE: WordPerfect (Version 5.1)  
 / CURRENT APPLICATION DATA: 08/471,613  
 / FILING DATE: 06-June-1995  
 / CLASSIFICATION: 435  
 / NAME: Paul T. Clark 30, 152  
 / ADDRESS: 225 Franklin Street  
 / TELEPHONE: (617) 542-5070  
 / TELECOMMUNICATION INFORMATION:  
 / TELEFAX: 200134 542-8906  
 / INFORMATION FOR SEQ ID NO: 1:

CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul T. Clark 30, 162  
REFERENCE/DOCKET NUMBER: 007966/229901  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8966  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 406  
STRAND: acid  
TOPOLOGY: N/A  
US-08-471-613-1  
Query Match 78.7%; Score 1665; DB 2; Length 406;  
Best Local Similarity 78.5%; Pred. No. 5,98-148;  
Matches 328; Conservative 38; Mismatches 40; Indels 12; Gaps 2;  
CY 1 MALVGEELKAADEVEFVNCASAFKPTGCGRLAHYLSHNEHQLLDKNTNRNMLFAOR 60  
DB 1 MALDGEQMLEEE-----KAGSGDYLYLSKIEELQYLDKNSQNLMLDQCR 49  
CY 61 NDLNSRYMPEELDELOEGSVGEVYVWKGCKVLWVHEGKGVVDDIDKNDITXIT 120  
DB 60 NEANAKVRLRESDLOEGSVGEVYVWKGCKVLWVHEGKGVVDDIDKNDITXIT 109  
CY 121 PSTVALNDSVVLHLVPSKVDLWLVKVEYVDSITVMDIGALDQCKEIKVYELP 180  
DB 110 PNCVVALNDSVTLHLKILNVDVPSVWVWVNDKRVLLVHGEKGVVDDIDKNDIT 109  
CY 181 KAPRLTSLDQKQKVLVGPPTGKTLAAVVAHHTDCTFRVSGSELVQKTYGSSR 240  
DB 170 KPELFEALQKQKVLVGPPTGKTLAAVVAHHTDCTFRVSGSELVQKTYGSSR 228  
CY 241 MYRELTVAAEHAASITFMDIDISGSRRSGSSGSGVQKQLELNDQGFELASK 300  
DB 240 MYRELTVAAEHAASITFMDIDISGSRRSGSSGSGVQKQLELNDQGFELASK 300  
CY 301 KYTAAWTRIDDLAALPGRIDPKIFPPHRESRLDITKTSRPNVLMGIDITKTA 360  
DB 300 MYRELTVAAEHAASITFMDIDISGSRRSGSSGSGVQKQLELNDQGFELASK 300  
CY 289 IYVIMATNRIDMLSLALPGRIDPKIFPPHRESRLDITKTSRPNVLMGIDITKTA 348  
DB 361 ERMNGSGAEELVYCTFCKEALRRVYVTCDFEYVAVVWVWVWVWVWVWVWVWVW 418  
CY 349 ELMPASGAEVQCTFCKEALRRVYVTCDFEYVAVVWVWVWVWVWVWVWVWVW 406  
RESULT 4  
Sequence 1, Application PC/TUS9310443  
GENERAL INFORMATION:  
APPLICANT: Jee W. Lee  
TITLE OF INVENTION: NUCLEIC HORMONE RECEPTOR-  
INTERACTING POLYPEPTIDES AND  
TITLE OF INVENTION: RELATED MOLECULES AND METHODS  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Jee W. Richardson  
ADDRESS: 25 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
OPERATING SYSTEM: MS-DOS 5.02 or 5.25  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/10443  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA: 07/969, 116  
FILING DATE: October 30, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul T. Clark 30, 162  
REFERENCE/DOCKET NUMBER: 007966/099002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8966  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 406  
STRAND: acid  
TOPOLOGY: N/A  
US-08-471-613-1  
Query Match 78.7%; Score 1665; DB 5; Length 406;  
Best Local Similarity 78.5%; Pred. No. 5,98-148;  
Matches 328; Conservative 38; Mismatches 40; Indels 12; Gaps 2;  
CY 1 MALVGEELKAADEVEFVNCASAFKPTGCGRLAHYLSHNEHQLLDKNTNRNMLFAOR 60  
DB 1 MALDGEQMLEEE-----KAGSGDYLYLSKIEELQYLDKNSQNLMLDQCR 49  
CY 61 NDLNSRYMPEELDELOEGSVGEVYVWKGCKVLWVHEGKGVVDDIDKNDITXIT 120  
DB 60 NEANAKVRLRESDLOEGSVGEVYVWKGCKVLWVHEGKGVVDDIDKNDITXIT 109  
CY 121 PSTVALNDSVVLHLVPSKVDLWLVKVEYVDSITVMDIGALDQCKEIKVYELP 180  
DB 110 PNCVVALNDSVTLHLKILNVDVPSVWVWVNDKRVLLVHGEKGVVDDIDKNDIT 109  
CY 181 KAPRLTSLDQKQKVLVGPPTGKTLAAVVAHHTDCTFRVSGSELVQKTYGSSR 240  
DB 170 KPELFEALQKQKVLVGPPTGKTLAAVVAHHTDCTFRVSGSELVQKTYGSSR 228  
CY 241 MYRELTVAAEHAASITFMDIDISGSRRSGSSGSGVQKQLELNDQGFELASK 300  
DB 240 MYRELTVAAEHAASITFMDIDISGSRRSGSSGSGVQKQLELNDQGFELASK 300  
CY 301 KYTAAWTRIDDLAALPGRIDPKIFPPHRESRLDITKTSRPNVLMGIDITKTA 360  
DB 300 MYRELTVAAEHAASITFMDIDISGSRRSGSSGSGVQKQLELNDQGFELASK 300  
CY 289 IYVIMATNRIDMLSLALPGRIDPKIFPPHRESRLDITKTSRPNVLMGIDITKTA 348  
DB 361 ERMNGSGAEELVYCTFCKEALRRVYVTCDFEYVAVVWVWVWVWVWVWVWVWVW 418  
CY 349 ELMPASGAEVQCTFCKEALRRVYVTCDFEYVAVVWVWVWVWVWVWVWVWVW 406  
RESULT 5  
US-08-222-719-2  
Sequence 2, Application US/08222719  
GENERAL INFORMATION:  
APPLICANT: David D. Moore  
TITLE OF INVENTION: NUCLEIC HORMONE RECEPTOR-  
INTERACTING POLYPEPTIDES AND  
TITLE OF INVENTION: RELATED MOLECULES AND METHODS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSES: David D. Richardson  
ADDRESS: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02110-2804



```

US-08-471-613-2
/ Sequence 2, Application US/08471613
/ Patent No. 5962256
/ GENERAL INFORMATION:
/ INVENTOR: J. W. Lee
/ APPLICANT: J. W. Lee
/ TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR-
/ INTERACTING POLYPEPTIDES AND
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: Fish & Richardson P.C.
/ City: Boston
/ State: Massachusetts
/ COUNTRY: U.S.A.
/ ZIP CODE: 02108
/ CONTACT PERSON: J. W. Lee
/ MEDIAN TYPE: 3.5" Diskette, 1.44 MB
/ COMPUTER: IBM PS/2 Model 502 or 555X
/ OPERATING SYSTEM: MS-DOS (Version 5.0)
/ SOFTWARE: Microsoft (Version 5.1)
/ CURRENT APPLICATION DATA:
/ FILING DATE: 06-June-1995
/ PRIORITY DATE: 06-June-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/222,719
/ FILING DATE: 04-April-1994
/ PRIORITY DATE: 04-April-1994
/ APPLICATION NUMBER: 07/969,136
/ FILING DATE: October 30, 1992
/ CLASSIFICATION: 418
/ APPLICATION INFORMATION:
/ NAME: Paul T. Clark
/ REGISTRATION NUMBER: 30,162
/ REFERENCE: POKET NUMBER: 07/66/229001
/ TELEPHONE: (617) 542-5070
/ TELEFAX: (617) 542-8906
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 405 acid
/ STRANDEDNESS:
/ TOPOLOGY: N/A
/ US-08-471-613-2
Query Match
Best Local Similarity 70.81; Score 1497.5; DB 2; Length 405;
Matches 289; Conservative 49; Mismatches 51; Indels 1; Gaps 1;
30 GAAHYSAIHQDIAHJQKTHMLRLEKQNDNSRYWALRESLQDLOFGSYGVGV 89
18 GIKPEYFKQVGTETKIKSTENGSRLEKQNALMDKFRFKQERLQDEFSYGVGV 77
90 WAGKIVLVYVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 149
76 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 137
150 KRWYVPGSTVDIGDQDQIKETKEVLEPIPIPELPSLGIQKQVLYVGPQCGIT 209
138 MEKVPDSTVDIGDQDQIKETKEVLEPIPIPELPSLGIQKQVLYVGPQCGIT 197
78 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 137
130 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 149
118 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 137
196 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 269
270 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 257
258 VEGSGSDSEVQGVNMLQDLPFVSNQIKIIMVNRDLQDLPALLRQKIDMKIIP 116

```

```

330 PPRNSRDLTIKHSRPNVWVGQIDKATFVNGSGRLEKATCTENGSLFRFNV 389
317 PPSVAAMALRILSHSRNNTGIDILNRKVRKNGKSGDQVGVCTEXQVABRSH 376
390 VYSGAPRNVNVAHSHSRNNTGIDILNRKVRKNGKSGDQVGVCTEXQVABRSH 418
377 VTGDPELVAGVWNNQDITLSVAKLEK 405
RESULT 8
PCT-US93-10443-2
/ Sequence 2, Application PCT/US9310443
/ GENERAL INFORMATION:
/ INVENTOR: J. W. Lee
/ APPLICANT: J. W. Lee
/ TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR-
/ INTERACTING POLYPEPTIDES AND
/ NUMBER OF SEQUENCES: 30
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: Fish & Richardson
/ City: Boston
/ State: Massachusetts
/ COUNTRY: U.S.A.
/ ZIP CODE: 02108
/ CONTACT PERSON: J. W. Lee
/ MEDIAN TYPE: 3.5" Diskette, 1.44 MB
/ COMPUTER: IBM PS/2 Model 502 or 555X
/ OPERATING SYSTEM: MS-DOS (Version 5.0)
/ SOFTWARE: Microsoft (Version 5.1)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US93/10443
/ CLASSIFICATION:
/ APPLICATION INFORMATION:
/ NAME: Paul T. Clark
/ REGISTRATION NUMBER: 07/969,136
/ REFERENCE: POKET NUMBER: 07/66/229001
/ TELEPHONE: (617) 542-5070
/ TELEFAX: (617) 542-8906
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 405 acid
/ STRANDEDNESS:
/ TOPOLOGY: N/A
/ PCT-US93-10443-2
Query Match
Best Local Similarity 70.81; Score 1497.5; DB 5; Length 405;
Matches 289; Conservative 49; Mismatches 51; Indels 1; Gaps 1;
30 GAAHYSAIHQDIAHJQKTHMLRLEKQNDNSRYWALRESLQDLOFGSYGVGV 89
18 GIKPEYFKQVGTETKIKSTENGSRLEKQNALMDKFRFKQERLQDEFSYGVGV 77
90 WAGKIVLVYVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 149
76 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 137
150 KRWYVPGSTVDIGDQDQIKETKEVLEPIPIPELPSLGIQKQVLYVGPQCGIT 209
138 MEKVPDSTVDIGDQDQIKETKEVLEPIPIPELPSLGIQKQVLYVGPQCGIT 197
78 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 137
130 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 149
118 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 137
196 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 269
270 IYSDKIVLVVNGPKVYVDAVDIDITKTSRYVALPNDGVVLYLVLPKQDPLVIM 257
198 LAAVAHPTCKCFRVSQALVQVYKIGSRNWELEPVNARSHASITKMOETISGSTR 257

```









Db 367 PMKXRNKADSKLESJLDYK 367

Search completed: November 9, 2003, 12:49:36  
Cdp time : 29 secs

---



QY 206 GKTLLAAVHNTDCTPIRVSSEGLVOKYIGESGHWELFVMAEBAHPS11PMOD1DS1 265  
 Db 191 GKTLLAAVHNTDCTPIRVSSEGLVOKYIGESGHWELFVMAEBAHPS11PMOD1DS1 260  
 QY 266 GSARSSGSSGDSDEQGTWLELLQDOFEPAANKYKVLAAWNR1D1DQALLPGRIDR 325  
 Db 251 GSARSSGSSGDSDEQGTWLELLQDOFEPAANKYKVLAAWNR1D1DQALLPGRIDR 309  
 QY 326 KIEFPFMESELD1K1KHRSRNNVAMGDI2KJLAEONQAGSGLVAVCEGOWFALE 385  
 Db 310 KIEFPFMESELD1K1KHRSRNNVAMGDI2KJLAEONQAGSGLVAVCEGOWFALE 369  
 QY 366 RRAYVQDEPMAVAAVWKKETSRNCEKRIK 418  
 Db 370 RRAYVQDEPMAVAAVWKKETSRNCEKRIK 402

RESULT 2  
 US-10-128-714-3369  
 / Publication No. US/0032565  
 / Publication No. US20030180993A1  
 / GENERAL INFORMATION  
 / APPLICANT: Teriy, Reemer D.  
 / APPLICANT: Hui, Weng  
 / APPLICANT: Chao, Bing Boone  
 / APPLICANT: Charles, Busee  
 / APPLICANT: Howard, Busee  
 / TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery  
 / CURRENT APPLICATION NUMBER: US/01/032,565  
 / CURRENT FILING DATE: 2001-12-20  
 / NUMBER OF SEQ ID NOS: 8000  
 / SEQ ID NO: 325  
 / SEQ ID NO: 326  
 / SEQ ID NO: 327  
 / LENGTH: 401  
 / TYPE: PRT  
 / FUNCTION: candida albicans  
 US-10-128-714-3369

Query Match 72.8%; Score 1540.5; DB 12; Length 401;  
 Best Local Similarity 76.3%; Pred. No. 4,18-133;  
 Matches 287; Conservative 49; Mismatches 42; Indels 1; Caps 1;  
 QY 30 GRLHYSLSNTHRHQLLKKQTNHMLRLEQRNDLSNRNMLREELDLQEPGSGVGEVKK 69  
 Db 14 GRLHYSLSNTHRHQLLKKQTNHMLRLEQRNDLSNRNMLREELDLQEPGSGVGEVKK 73  
 QY 90 VNGKNNVLYVWPEKXVYLDKND1DK1TSTGRVALRANQKRLKQELKLEDSGVGEVKK 143  
 Db 74 VNGKNNVLYVWPEKXVYLDKND1DK1TSTGRVALRANQKRLKQELKLEDSGVGEVKK 139  
 QY 150 KYEVPDSTYDMLQDOQIKETKEVLEL1HPELFPESLGIAGQPKVLYLGGPGTCKTL 209  
 Db 134 KYEVPDSTYDMLQDOQIKETKEVLEL1HPELFPESLGIAGQPKVLYLGGPGTCKTL 193  
 QY 144 MVEVPSGLDVGQGLVOKYIGESGHWELFVMAEBAHPS11PMOD1DS1GQATL 265  
 Db 210 MVEVPSGLDVGQGLVOKYIGESGHWELFVMAEBAHPS11PMOD1DS1GQATL 260  
 QY 270 MESSGNDSEQGTWLELLQDOFEPAANKYKVLAAWNR1D1DQALLPGRIDR 325  
 Db 194 MESSGNDSEQGTWLELLQDOFEPAANKYKVLAAWNR1D1DQALLPGRIDR 309  
 QY 284 VE-GSSGDSGSEVQGTWLELLQDOFEPAANKYKVLAAWNR1D1DQALLPGRIDR 312  
 Db 330 PPRMSESLD1K1KHRSRNNVAMGDI2KJLAEONQAGSGLVAVCEGOWFALE 385  
 QY 310 PPRMSESLD1K1KHRSRNNVAMGDI2KJLAEONQAGSGLVAVCEGOWFALE 369  
 QY 390 YQDEPMAVAAVWKKETSRNCEKRIK 418  
 Db 373 YQDEPMAVAAVWKKETSRNCEKRIK 401

/ Sequence 3369, Application US/0128714  
 / Publication No. US20030119013A1  
 / GENERAL INFORMATION  
 / APPLICANT: Hui, Weng  
 / APPLICANT: Teriy, Reemer D.  
 / APPLICANT: Chao, Bing Boone  
 / APPLICANT: Charles, Busee  
 / APPLICANT: Howard, Busee  
 / TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
 / TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
 / FILE NUMBER: 10182-018-999  
 / CURRENT FILING DATE: 2002-04-23  
 / CURRENT FILING DATE: 2002-04-23  
 / PRIOR APPLICATION NUMBER: US 60/285,697  
 / PRIOR APPLICATION NUMBER: US 60/285,697  
 / PRIOR APPLICATION NUMBER: US 60/287,066  
 / PRIOR APPLICATION NUMBER: US 60/287,066  
 / PRIOR FILING DATE: 2001-06-05  
 / PRIOR FILING DATE: 2001-06-05  
 / PRIOR APPLICATION NUMBER: US 60/303,899  
 / PRIOR APPLICATION NUMBER: US 60/303,899  
 / PRIOR FILING DATE: 2001-08-31  
 / PRIOR FILING DATE: 2001-08-31  
 / NUMBER OF SEQ ID NOS: 8603  
 / SEQ ID NO: 389  
 / SEQ ID NO: 390  
 / LENGTH: 389  
 / TYPE: PRT  
 / FUNCTION: Aspergillus fumigatus  
 US-10-128-714-3369

Query Match 71.3%; Score 1508.5; DB 15; Length 389;  
 Best Local Similarity 77.8%; Pred. No. 3,46-130;  
 Matches 304; Conservative 34; Mismatches 51; Indels 1; Caps 1;  
 QY 31 LRLHYSLSNTHRHQLLKKQTNHMLRLEQRNDLSNRNMLREELDLQEPGSGVGEVKK 90  
 Db 3 LRLHYSLSNTHRHQLLKKQTNHMLRLEQRNDLSNRNMLREELDLQEPGSGVGEVKK 90  
 QY 91 MGRKNNVLYVWPEKXVYLDKND1DK1TSTGRVALRANQKRLKQELKLEDSGVGEVKK 150  
 Db 63 MGRKNNVLYVWPEKXVYLDKND1DK1TSTGRVALRANQKRLKQELKLEDSGVGEVKK 122  
 QY 151 VNGKNNVLYVWPEKXVYLDKND1DK1TSTGRVALRANQKRLKQELKLEDSGVGEVKK 210  
 Db 123 VNGKNNVLYVWPEKXVYLDKND1DK1TSTGRVALRANQKRLKQELKLEDSGVGEVKK 182  
 QY 211 KYEVPDSTYDMLQDOQIKETKEVLEL1HPELFPESLGIAGQPKVLYLGGPGTCKTL 270  
 Db 183 KYEVPDSTYDMLQDOQIKETKEVLEL1HPELFPESLGIAGQPKVLYLGGPGTCKTL 242  
 QY 271 MESSGNDSEQGTWLELLQDOFEPAANKYKVLAAWNR1D1DQALLPGRIDR 325  
 Db 243 MESSGNDSEQGTWLELLQDOFEPAANKYKVLAAWNR1D1DQALLPGRIDR 301  
 QY 331 PPRMSESLD1K1KHRSRNNVAMGDI2KJLAEONQAGSGLVAVCEGOWFALE 385  
 Db 302 PPRMSESLD1K1KHRSRNNVAMGDI2KJLAEONQAGSGLVAVCEGOWFALE 361  
 QY 391 YQDEPMAVAAVWKKETSRNCEKRIK 418  
 Db 362 YQDEPMAVAAVWKKETSRNCEKRIK 389

RESULT 4  
 US-10-128-714-8369  
 / Sequence 8369, Application US/0128714  
 / Publication No. US20030119013A1  
 / GENERAL INFORMATION  
 / APPLICANT: Hui, Weng  
 / APPLICANT: Teriy, Reemer D.  
 / APPLICANT: Chao, Bing Boone  
 / APPLICANT: Charles, Busee  
 / APPLICANT: Howard, Busee  
 / TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
 / TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
 / FILE NUMBER: 10182-018-999  
 / CURRENT FILING DATE: 2002-04-23  
 / CURRENT FILING DATE: 2002-04-23  
 / PRIOR APPLICATION NUMBER: US 60/285,697  
 / PRIOR APPLICATION NUMBER: US 60/285,697  
 / PRIOR APPLICATION NUMBER: US 60/287,066  
 / PRIOR APPLICATION NUMBER: US 60/287,066  
 / PRIOR FILING DATE: 2001-06-05  
 / PRIOR FILING DATE: 2001-06-05  
 / PRIOR APPLICATION NUMBER: US 60/303,899  
 / PRIOR APPLICATION NUMBER: US 60/303,899  
 / PRIOR FILING DATE: 2001-08-31  
 / PRIOR FILING DATE: 2001-08-31  
 / NUMBER OF SEQ ID NOS: 8603  
 / SEQ ID NO: 389  
 / SEQ ID NO: 390  
 / LENGTH: 389  
 / TYPE: PRT  
 / FUNCTION: Aspergillus fumigatus  
 US-10-128-714-8369













```

Db      39  LRIMGEYOLTHEONTHREVERENDOKKIMHROLPYLVONVELLDIDYEAAREONV 98
Ov      92  -----GNKVLVWHEBKYVVDIDKIDITKTPTWLANDSYVLAHVASRVD 143
Db      99  IDLONRWGKSNV-ITSTRQITVYPLIGLVWHEKFKRODLGVKDSVILDTLPAED 157
Ov      144  PLVIMKEVPESTTOMIGELDOQIKETVELPIHPELFTESLGIQKQVULYGP 203
Db      156  NRYAMVEDKFTKNTDIDKQIDELVENVWKEKEEFTYIGIAPKSNVYGP 217
Ov      204  GTGKTLARVAAHHTCTFINSSSELYQYTGGSRWVEELFWAEEAPSIITWDEID 263
Db      218  GYKNTLARKKASANTLAKMGQVQWVIGDQKAPKQKAKENRPSIITDSD 277
Ov      264  SIGSARWSSGNGDSEWNTMELINQDGEASNKIKLAKTNRIDIDQALLPGR 323
Db      278  AVOTKIDFS-EXSDNFRQKTHSLQDQSDQSDRINKLANKNIDVDYALSGEL 336
Ov      324  DKRIEPTPEESHLDILKHSRNNLMQCIDLKKIAPQKNGSGALHAYCTEKNFAL 383
Db      337  DKRIEPTPEESHENKLDQIHSGKANDANMAHESSTDEFGQKQKNGVGNKNTAL 396
Ov      384  RBRVYVTODEFENAVAKYMEKETERNSL 413
Db      397  KKNSTVGHENTVAMALVQKSKOTNGI 426

```

Search completed: November 9, 2003, 12:57:17  
 Job time : 80 sec0



```

RESULT 2
A:Accession: 1-406 <K1>
A:Species: 26S proteasome chain p45 - human
C:Update: 19-Mar-1997. Sequence revision 05-May-1997. Rtext_change 19-Jan-2001
C:Accession: S65363; S65488; I55510
Pfam loci: 363, 151-156, 195
A:Title: cDN: Cloning of a new putative ATPase subunit p45 of the human 26S proteasome,
A:Reference number: I55510; MIMD:95246863; PMID:7729537
A:Molecule type: mRNA
A:Residues: 1-406 <K1>
A:Accession: EMBL:U94467; NID:9976226; PIDN:BA007919.1; PID:9976227
A:Species: 26S proteasome chain p45 - human
C:Update: 06-Sep-1996. Sequence revision 27-Feb-1997. Rtext_change 19-Jan-2001
C:Accession: S60343
Pfam loci: 363, 151-156, 195
A:Title: Interaction of thymidylate synthase with a conserved transcriptional mediator
A:Reference number: S60343; MIMD:95174891; PMID:7670182
A:Accession: S60343
A:Species: 26S proteasome chain p45 - human
A:Molecule type: mRNA
Query Match 78 2% Score 165; DB 2; Length 406;
Similarity 1; P100; Region: nucleotide-binding motif A (P-loop)
Matches 330; Conservative 37; Mismatches 39; Indels 12; Gaps 2;
1 MALVGEELKAAEGEVENKSNKAFPTGSGSARHTYSLEINHHNLLQKNTKNNLLEGR 60
1 MALDGEVMELEED-----KASGLQVYLSIELEQLVINDSSNLRDQR 49
61 NDLSRVMRLREELQLODEPSVGEVYVWAKKLVVHFESEKVVINDINDITKIT 120
50 NEDAKNVLRLREELQLODEPSVGEVYVWAKKLVVHFESEKVVINDINDITKIT 109
121 PSTRYALNLSVVLNVLVPEKVPVYVWAKKLVVHFESEKVVINDINDITKIT 180
120 PSTRYALNLSVVLNVLVPEKVPVYVWAKKLVVHFESEKVVINDINDITKIT 169
170 KRPFLPSGLIDVPGKGLTVYPPGKGLTLAAVAARHDCFTIRVSGSELVYIGESGR 240
181 KRPFLPSGLIDVPGKGLTVYPPGKGLTLAAVAARHDCFTIRVSGSELVYIGESGR 229
170 KRPFLPSGLIDVPGKGLTVYPPGKGLTLAAVAARHDCFTIRVSGSELVYIGESGR 229
241 MYRELPLVAARHPRHAIIPMDIDISGASNGSGSGSGSGSGSGSGSGSGSGSGSG 300
230 MYRELPLVAARHPRHAIIPMDIDISGASNGSGSGSGSGSGSGSGSGSGSGSGSG 288
301 IYVLAATRRDIDLOALFALRPRDKLEPFPREESDIDLIHSMRPLDIDLIKIA 360
289 IYVLAATRRDIDLOALFALRPRDKLEPFPREESDIDLIHSMRPLDIDLIKIA 348
361 EKNMGSGALAVCTACGAPALPRRRVYVDEPDAVAAKKKEKEGMSLRLK 418
349 ELPDGSAGVAGCTACGAPALPRRRVYVDEPDAVAAKKKEGMSLRLK 406

```

```

A:Residues: 1-406 <L26>
A:Accession: GB:U136810; NID:9695369; PIDN:MA01735.1; PID:9695370
C:Superfamily: ATP-dependent 26S proteasome; Pfam/SC18/CD48-type ATP-binding domain
F:463-373/Domain: Pfam/SC18/CD48-type ATP-binding domain homology <VAP>
F:190-197/Region: nucleotide-binding motif A (P-loop)
Query Match 78 2% Score 166; DB 2; Length 406;
Similarity 1; P100; Region: nucleotide-binding motif A (P-loop)
Matches 338; Conservative 38; Mismatches 40; Indels 12; Gaps 2;
1 MALVGEELKAAEGEVENKSNKAFPTGSGSARHTYSLEINHHNLLQKNTKNNLLEGR 60
1 MALDGEVMELEED-----KASGLQVYLSIELEQLVINDSSNLRDQR 49
61 NDLSRVMRLREELQLODEPSVGEVYVWAKKLVVHFESEKVVINDINDITKIT 120
50 NEDAKNVLRLREELQLODEPSVGEVYVWAKKLVVHFESEKVVINDINDITKIT 109
121 PSTRYALNLSVVLNVLVPEKVPVYVWAKKLVVHFESEKVVINDINDITKIT 180
120 PSTRYALNLSVVLNVLVPEKVPVYVWAKKLVVHFESEKVVINDINDITKIT 169
170 KRPFLPSGLIDVPGKGLTVYPPGKGLTLAAVAARHDCFTIRVSGSELVYIGESGR 240
181 KRPFLPSGLIDVPGKGLTVYPPGKGLTLAAVAARHDCFTIRVSGSELVYIGESGR 229
170 KRPFLPSGLIDVPGKGLTVYPPGKGLTLAAVAARHDCFTIRVSGSELVYIGESGR 229
241 MYRELPLVAARHPRHAIIPMDIDISGASNGSGSGSGSGSGSGSGSGSGSGSGSG 300
230 MYRELPLVAARHPRHAIIPMDIDISGASNGSGSGSGSGSGSGSGSGSGSGSGSG 288
301 IYVLAATRRDIDLOALFALRPRDKLEPFPREESDIDLIHSMRPLDIDLIKIA 360
289 IYVLAATRRDIDLOALFALRPRDKLEPFPREESDIDLIHSMRPLDIDLIKIA 348
361 EKNMGSGALAVCTACGAPALPRRRVYVDEPDAVAAKKKEKEGMSLRLK 418
349 ELPDGSAGVAGCTACGAPALPRRRVYVDEPDAVAAKKKEGMSLRLK 406

```













GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd

ON protein - protein search, using

Run on: November 9, 2003, 11:17:06 ; Search time 24 Seconds  
(without placement)

Title: US-09-462-972-2  
Page: 0116

Scoring table: BLOSUM62

Searched: 127863 segs, 47026705 residues  
Total number of hits satisfying chosen parameters

Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0

Database : SwissProt\_41:

score greater than or equal to the score of the choice being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match Length	DB	Description
1	1,680	79.4	405	1	PR88_HUMAN
2	1,660.5	78.5	405	1	PR88_DROMA
3	1,655.5	78.2	402	1	PR88_MANS
4	1,655	78.2	402	1	PR88_MANS
5	1,530	77.0	398	1	PR88_DIC1
6	1,511	76.1	403	1	PR88_MARO
7	1,197	75.5	403	1	PR88_SOPHO
8	1,504.5	72.1	401	1	PR88_YEAST
9	1,497.5	45.3	396	1	PR88_PRRU
10	957.5	45.3	396	1	PR88_PRRU
11	955.5	45.2	399	1	PR88_PRRU
12	937.5	44.3	399	1	PR88_PRRU
13	929	43.9	398	1	PR88_ACHU
14	928	43.9	398	1	PR88_ACHU
15	906	42.8	400	1	PR88_MER1
16	903	42.7	443	1	PR84_CIEB
17	901	42.6	440	1	PR87_HUMAN
18	895	42.3	459	1	PR81_CIEB
19	895	42.3	459	1	PR81_CIEB
20	894.5	42.3	459	1	PR87_PRRU
21	893.5	42.3	456	1	PR87_PRRU
22	893	42.2	456	1	PR87_PRRU
23	888	42.0	433	1	PR87_HUMAN
24	888	42.0	433	1	PR87_HUMAN
25	888	42.0	433	1	PR87_HUMAN
26	888	42.0	433	1	PR87_HUMAN
27	880	41.6	433	1	PR87_PRRU
28	880	41.6	433	1	PR87_PRRU
29	878.5	41.5	445	1	PR87_CIEB
30	864	40.8	406	1	PR87_CIEB
31	858	40.6	448	1	PR87_CIEB
32	853	40.3	448	1	PR87_CIEB
33	853	40.3	448	1	PR87_CIEB
34	853	40.3	448	1	PR87_CIEB
35	853	40.3	448	1	PR87_CIEB
36	853	40.3	448	1	PR87_CIEB
37	853	40.3	448	1	PR87_CIEB
38	853	40.3	448	1	PR87_CIEB
39	853	40.3	448	1	PR87_CIEB
40	853	40.3	448	1	PR87_CIEB
41	853	40.3	448	1	PR87_CIEB
42	853	40.3	448	1	PR87_CIEB
43	853	40.3	448	1	PR87_CIEB
44	853	40.3	448	1	PR87_CIEB
45	853	40.3	448	1	PR87_CIEB
46	853	40.3	448	1	PR87_CIEB
47	853	40.3	448	1	PR87_CIEB
48	853	40.3	448	1	PR87_CIEB
49	853	40.3	448	1	PR87_CIEB
50	853	40.3	448	1	PR87_CIEB
51	853	40.3	448	1	PR87_CIEB
52	853	40.3	448	1	PR87_CIEB
53	853	40.3	448	1	PR87_CIEB
54	853	40.3	448	1	PR87_CIEB
55	853	40.3	448	1	PR87_CIEB
56	853	40.3	448	1	PR87_CIEB
57	853	40.3	448	1	PR87_CIEB
58	853	40.3	448	1	PR87_CIEB
59	853	40.3	448	1	PR87_CIEB
60	853	40.3	448	1	PR87_CIEB
61	853	40.3	448	1	PR87_CIEB
62	853	40.3	448	1	PR87_CIEB
63	853	40.3	448	1	PR87_CIEB
64	853	40.3	448	1	PR87_CIEB
65	853	40.3	448	1	PR87_CIEB
66	853	40.3	448	1	PR87_CIEB
67	853	40.3	448	1	PR87_CIEB
68	853	40.3	448	1	PR87_CIEB
69	853	40.3	448	1	PR87_CIEB
70	853	40.3	448	1	PR87_CIEB
71	853	40.3	448	1	PR87_CIEB
72	853	40.3	448	1	PR87_CIEB
73	853	40.3	448	1	PR87_CIEB
74	853	40.3	448	1	PR87_CIEB
75	853	40.3	448	1	PR87_CIEB
76	853	40.3	448	1	PR87_CIEB
77	853	40.3	448	1	PR87_CIEB
78	853	40.3	448	1	PR87_CIEB
79	853	40.3	448	1	PR87_CIEB
80	853	40.3	448	1	PR87_CIEB
81	853	40.3	448	1	PR87_CIEB
82	853	40.3	448	1	PR87_CIEB
83	853	40.3	448	1	PR87_CIEB
84	853	40.3	448	1	PR87_CIEB
85	853	40.3	448	1	PR87_CIEB
86	853	40.3	448	1	PR87_CIEB
87	853	40.3	448	1	PR87_CIEB
88	853	40.3	448	1	PR87_CIEB
89	853	40.3	448	1	PR87_CIEB
90	853	40.3	448	1	PR87_CIEB
91	853	40.3	448	1	PR87_CIEB
92	853	40.3	448	1	PR87_CIEB
93	853	40.3	448	1	PR87_CIEB
94	853	40.3	448	1	PR87_CIEB
95	853	40.3	448	1	PR87_CIEB
96	853	40.3	448	1	PR87_CIEB
97	853	40.3	448	1	PR87_CIEB
98	853	40.3	448	1	PR87_CIEB
99	853	40.3	448	1	PR87_CIEB
100	853	40.3	448	1	PR87_CIEB

34	850	40.2	420	1	P8NF_METAC	Q8Y188 methanarsarc
35	850	40.2	421	1	P8NF_METAC	Q8Y188 methanarsarc
36	849.5	40.1	467	1	P8NF_YEAST	P1339 saccharomy
37	849.5	40.1	467	1	P8NF_YEAST	P1339 saccharomy
38	813	38.6	409	1	P8NF_AE09P	G09247 aeropyrum
39	816.5	38.6	437	1	SUGO_YEAST	P15339 saccharom
40	816.5	38.6	439	1	P8NF_HUMAN	P17940 homo sapien
41	815	38.6	439	1	P8NF_YEAST	Q8J359 lactuca nov
42	815	38.6	439	1	P8NF_YEAST	Q8J359 lactuca nov
43	807.5	38.2	442	1	P8NF_MOUSE	O08655 mus musculu
44	808.5	38.2	404	1	P8NF_XENLA	Q04236 xenopus lae
45	807.5	38.2	392	1	P8NF_SULTO	Q07502 sulfolobus

## ALIGNMENT

Accession	Source	Standard	Ref.	406 Mb.
AF000001	Human	Standard	1	
AF000002	Human	Standard	2	
AF000003	Human	Standard	3	
AF000004	Human	Standard	4	
AF000005	Human	Standard	5	
AF000006	Human	Standard	6	
AF000007	Human	Standard	7	
AF000008	Human	Standard	8	
AF000009	Human	Standard	9	
AF000010	Human	Standard	10	
AF000011	Human	Standard	11	
AF000012	Human	Standard	12	
AF000013	Human	Standard	13	
AF000014	Human	Standard	14	
AF000015	Human	Standard	15	
AF000016	Human	Standard	16	
AF000017	Human	Standard	17	
AF000018	Human	Standard	18	
AF000019	Human	Standard	19	
AF000020	Human	Standard	20	
AF000021	Human	Standard	21	
AF000022	Human	Standard	22	
AF000023	Human	Standard	23	
AF000024	Human	Standard	24	
AF000025	Human	Standard	25	
AF000026	Human	Standard	26	
AF000027	Human	Standard	27	
AF000028	Human	Standard	28	
AF000029	Human	Standard	29	
AF000030	Human	Standard	30	
AF000031	Human	Standard	31	
AF000032	Human	Standard	32	
AF000033	Human	Standard	33	
AF000034	Human	Standard	34	
AF000035	Human	Standard	35	
AF000036	Human	Standard	36	
AF000037	Human	Standard	37	
AF000038	Human	Standard	38	
AF000039	Human	Standard	39	
AF000040	Human	Standard	40	
AF000041	Human	Standard	41	
AF000042	Human	Standard	42	
AF000043	Human	Standard	43	
AF000044	Human	Standard	44	
AF000045	Human	Standard	45	
AF000046	Human	Standard	46	
AF000047	Human	Standard	47	
AF000048	Human	Standard	48	
AF000049	Human	Standard	49	
AF000050	Human	Standard	50	
AF000051	Human	Standard	51	
AF000052	Human	Standard	52	
AF000053	Human	Standard	53	
AF000054	Human	Standard	54	
AF000055	Human	Standard	55	
AF000056	Human	Standard	56	
AF000057	Human	Standard	57	
AF000058	Human	Standard	58	
AF000059	Human	Standard	59	
AF000060	Human	Standard	60	
AF000061	Human	Standard	61	
AF000062	Human	Standard	62	
AF000063	Human	Standard	63	
AF000064	Human	Standard	64	
AF000065	Human	Standard	65	
AF000066	Human	Standard	66	
AF000067	Human	Standard	67	
AF000068	Human	Standard	68	
AF000069	Human	Standard	69	
AF000070	Human	Standard	70	
AF000071	Human	Standard	71	
AF000072	Human	Standard	72	
AF000073	Human	Standard	73	
AF000074	Human	Standard	74	
AF000075	Human	Standard	75	
AF000076	Human	Standard	76	
AF000077	Human	Standard	77	
AF000078	Human	Standard	78	
AF000079	Human	Standard	79	
AF000080	Human	Standard	80	
AF000081	Human	Standard	81	
AF000082	Human	Standard	82	
AF000083	Human	Standard	83	
AF000084	Human	Standard	84	
AF000085	Human	Standard	85	
AF000086	Human	Standard	86	
AF000087	Human	Standard	87	
AF000088	Human	Standard	88	
AF000089	Human	Standard	89	
AF000090	Human	Standard	90	
AF000091	Human	Standard	91	
AF000092	Human	Standard	92	
AF000093	Human	Standard	93	
AF000094	Human	Standard	94	
AF000095	Human	Standard	95	
AF000096	Human	Standard	96	
AF000097	Human	Standard	97	
AF000098	Human	Standard	98	
AF000099	Human	Standard	99	
AF000100	Human	Standard	100	























```

50 SEQUENCE 430 AA: 46689 MW: 3FD294A6BD4A3DD CRG64;
Query Match 42.8%; Score 906; DB 1; Length 430;
Best Local Similarity 45.0%; Evid. No. 1.1e-50;
Matches 186; Conservative 86; Mismatches 133; Indels 8; Gaps 2;

4 VGVETKHAADVPVANCAPPTQCEGLNHYSLNTHHQLLRCKTHNLNRLEQNRNL 63
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 7 ISTEKKERKAFTEPEKEKSIINDSNLNDLKEELAEARINLESRIKLLEKEKL 66
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 64 N-----SIVMLFEELQLQPESSVGEVYVWAGNKLVLVYHNGGYVDIDCKYDI 116
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 67 ERENQJMKENSELLEFELDNRKVPPLVGVTVDKVSKRYVKSSTGSEFLVNSHFTVP 126
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 117 TKITSTVALNDSYVLLVLPSPVDPVLNWKYKVPSTYDMGGDDQKXIKETVI 176
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 127 DDLAQRVCLNQGLTVQVLPENKDYKAMVEDERPVRYEDIGLEKCKQREIVV 186
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 177 ELPIYHFEFLDIADQKQVLYGPGTCKLLAAVAHNDCTIRVSGSELVQYTG 236
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 187 ELTLHPELFEVQGIYGPPTGKTLAAVAATEVATIRVQSELVQYTG 246
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 237 EGSRVWELFVWABEAFSIIIPNDISIGSNWEGSGNPSQGTVALDLNQLDGE 296
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 247 EGASLVQDIFLAKEKASIIIFIDIDALANKRTDLTG-GREVOETLKQLLAMQGH 305
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 297 ASNKIKVLTATNRIDILQALLPQRIIDKXIEPPTNESEHDIKTHSRMYLNGIDL 356
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 306 AKQDVKIGATNRPDLIDPALIRGRPDRIIVRAPDEKGRSLIKIHTRQNLAEVNL 365
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 357 KXIAEKWNGSGALEIAYCTAKNGPALAPRYHTQDEPNAVAKMKETEK 409
QY : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 366 BELAQTGECVGAELKALCTEPMMAIRLELDVTVWDFKAVKIKMKKKVX 418
QY : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: November 9, 2003, 12:46:09  
 Job time : 25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Camgen Ltd.

OK protein - protein search, using sw model

Run on: November 9, 2003, 12:29:56; Search time 110 seconds

(without alignments)

US-09-462-972-2

1 MALVGEVHKAIEAEVFAKNC.....VAKVKKETKQNSLRMLK 418

Sequence: 2116

Scoring table:

Gapop 10.0, Gapext 0.5

830525 seqs; 25805604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length 0

Maximum DB seq length 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Display first 45 summaries

Database:

- 1: sp\_arctia;
- 2: sp\_bacteria;
- 3: sp\_bacteriophage;
- 4: sp\_bum;
- 5: sp\_invertebrate;
- 6: sp\_mammal;
- 7: sp\_mammal2;
- 8: sp\_mammal3;
- 9: sp\_mammal4;
- 10: sp\_mammal5;
- 11: sp\_mammal6;
- 12: sp\_mammal7;
- 13: sp\_mammal8;
- 14: sp\_mammal9;
- 15: sp\_mammal10;
- 16: sp\_mammal11;
- 17: sp\_mammal12;

Read No. is the number of results predicted by chance to have a score greater than the observed score. The score is the total score and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
1	1926	91.0	419	10	095C03		
2	1913	90.4	409	10	095E11		
3	1905.5	90.1	419	10	094B02		
4	1903.5	89.9	419	10	095E11		
5	1893.5	89.9	419	10	095E11		
6	1879.5	88.8	419	10	095E11		
7	1872.5	88.5	418	10	095C06		
8	1861	87.9	419	10	095C06		
9	1851	87.9	419	10	095C06		
10	1822.5	76.7	399	5	095E11		
11	1622.5	75.1	435	5	094E05		
12	1587.5	74.3	418	5	095E11		
13	1572	74.3	418	5	095E11		
14	1561	70.9	411	5	095E11		
15	1501	70.9	411	5	095E11		
16	1450	67.1	432	5	095E11		

17	1376	65.1	408	5	095C05		095C05 trypanosoma
18	1365.5	64.5	306	11	095K01		095K01 mus musculus
19	1327	62.7	348	11	095K02		095K02 mus musculus
20	1307.5	50.9	440	8	095E11		095E11 mus musculus
21	1270.5	50.9	440	8	095E11		095E11 mus musculus
22	922.5	44.1	217	5	095C09		095C09 amblyomma a
23	895	42.3	439	5	095C01		095C01 trypanosoma
24	888	42.0	435	11	095E04		095E04 mus musculus
25	887.5	41.5	390	5	095E11		095E11 mus musculus
26	877.5	41.5	390	5	095E11		095E11 mus musculus
27	877.5	41.5	390	5	095E11		095E11 mus musculus
28	866.5	41.1	355	5	095E11		095E11 mus musculus
29	866.5	41.1	355	5	095E11		095E11 mus musculus
30	866.5	41.0	355	5	095E11		095E11 mus musculus
31	867	41.0	389	11	095C09		095C09 mus musculus
32	865.5	40.9	421	10	095E11		095E11 mus musculus
33	865.5	40.7	397	5	095E11		095E11 mus musculus
34	865.5	40.7	397	5	095E11		095E11 mus musculus
35	860	40.6	393	5	095E01		095E01 trypanosoma
36	860	40.6	393	5	095E01		095E01 trypanosoma
37	858.5	40.6	443	10	095E11		095E11 mus musculus
38	858.5	40.6	443	10	095E11		095E11 mus musculus
39	858.5	40.6	443	10	095E11		095E11 mus musculus
40	857.5	40.5	396	5	095E11		095E11 mus musculus
41	856.5	40.5	396	5	095E11		095E11 mus musculus
42	856.5	40.5	396	5	095E11		095E11 mus musculus
43	856.5	40.5	396	5	095E11		095E11 mus musculus
44	855.5	40.4	459	3	095E11		095E11 mus musculus
45	855.5	40.4	443	10	095E11		095E11 mus musculus

# ALIGNMENTS

RESULT 1	PRELIMINARY	ERT	419 NA.
1	095C03		
2	095C03		
3	095C03		
4	095C03		
5	095C03		
6	095C03		
7	095C03		
8	095C03		
9	095C03		
10	095C03		
11	095C03		
12	095C03		
13	095C03		
14	095C03		
15	095C03		
16	095C03		
17	095C03		
18	095C03		
19	095C03		
20	095C03		
21	095C03		
22	095C03		
23	095C03		
24	095C03		
25	095C03		
26	095C03		
27	095C03		
28	095C03		
29	095C03		
30	095C03		
31	095C03		
32	095C03		
33	095C03		
34	095C03		
35	095C03		
36	095C03		
37	095C03		
38	095C03		
39	095C03		
40	095C03		
41	095C03		
42	095C03		
43	095C03		
44	095C03		
45	095C03		







5.1	Gaps	5	4	3	2	1	0
5.2	Attacks	4	3	2	1	0	0
5.3	Unbalanced	4	3	2	1	0	0
5.4	Unbalanced	4	3	2	1	0	0
5.5	Unbalanced	4	3	2	1	0	0
5.6	Unbalanced	4	3	2	1	0	0
5.7	Unbalanced	4	3	2	1	0	0
5.8	Unbalanced	4	3	2	1	0	0
5.9	Unbalanced	4	3	2	1	0	0
5.10	Unbalanced	4	3	2	1	0	0
5.11	Unbalanced	4	3	2	1	0	0
5.12	Unbalanced	4	3	2	1	0	0
5.13	Unbalanced	4	3	2	1	0	0
5.14	Unbalanced	4	3	2	1	0	0
5.15	Unbalanced	4	3	2	1	0	0
5.16	Unbalanced	4	3	2	1	0	0
5.17	Unbalanced	4	3	2	1	0	0
5.18	Unbalanced	4	3	2	1	0	0
5.19	Unbalanced	4	3	2	1	0	0
5.20	Unbalanced	4	3	2	1	0	0
5.21	Unbalanced	4	3	2	1	0	0
5.22	Unbalanced	4	3	2	1	0	0
5.23	Unbalanced	4	3	2	1	0	0
5.24	Unbalanced	4	3	2	1	0	0
5.25	Unbalanced	4	3	2	1	0	0
5.26	Unbalanced	4	3	2	1	0	0
5.27	Unbalanced	4	3	2	1	0	0
5.28	Unbalanced	4	3	2	1	0	0
5.29	Unbalanced	4	3	2	1	0	0
5.30	Unbalanced	4	3	2	1	0	0
5.31	Unbalanced	4	3	2	1	0	0
5.32	Unbalanced	4	3	2	1	0	0
5.33	Unbalanced	4	3	2	1	0	0
5.34	Unbalanced	4	3	2	1	0	0
5.35	Unbalanced	4	3	2	1	0	0
5.36	Unbalanced	4	3	2	1	0	0
5.37	Unbalanced	4	3	2	1	0	0
5.38	Unbalanced	4	3	2	1	0	0
5.39	Unbalanced	4	3	2	1	0	0
5.40	Unbalanced	4	3	2	1	0	0
5.41	Unbalanced	4	3	2	1	0	0
5.42	Unbalanced	4	3	2	1	0	0
5.43	Unbalanced	4	3	2	1	0	0
5.44	Unbalanced	4	3	2	1	0	0
5.45	Unbalanced	4	3	2	1	0	0
5.46	Unbalanced	4	3	2	1	0	0
5.47	Unbalanced	4	3	2	1	0	0
5.48	Unbalanced	4	3	2	1	0	0
5.49	Unbalanced	4	3	2	1	0	0
5.50	Unbalanced	4	3	2	1	0	0
5.51	Unbalanced	4	3	2	1	0	0
5.52	Unbalanced	4	3	2	1	0	0
5.53	Unbalanced	4	3	2	1	0	0
5.54	Unbalanced	4	3	2	1	0	0
5.55	Unbalanced	4	3	2	1	0	0
5.56	Unbalanced	4	3	2	1	0	0
5.57	Unbalanced	4	3	2	1	0	0
5.58	Unbalanced	4	3	2	1	0	0
5.59	Unbalanced	4	3	2	1	0	0
5.60	Unbalanced	4	3	2	1	0	0
5.61	Unbalanced	4	3	2	1	0	0
5.62	Un						







GN B1D4.170.  
 OS Neurospora crassa.  
 OC Neurospora crassa; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariata; Sordariaceae; Neurospora.  
 OX NCBI\_Taxid:5141;  
 RN SEQUENCE FROM N.A.  
 RA Schulte U, Alen V, Holsel J, Brandt P., Fartmann B., Holsel R.,  
 RA Nykter G., Mewes H.-W., Mannhap G.;  
 RA Submitted (May-2000) to the EMBL/Genbank/DBJ databases.  
 RL Submitted (May-2000) to the EMBL/Genbank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RC German Neurospora genome project;  
 RL Submitted (Nov-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL155528; GenBank: J05115.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR003593; AAA\_ATPase\_cent.  
 DR InterPro: IPR003590; AAA\_sub.  
 DR SMART: SM00383; AAA; 1.  
 DR TIGRFAMs: TIGR01242; 265945; 1.  
 DR PROSITE: PS00674; AAA; 1.  
 DR ATTBinding: protease.  
 SQ 5500000 309 AA; 43584 MW; C61397A10N793 CR664;  
 Query Match 72.2%; Score 1528.5; DB 3; Length 389; Gaps 1;  
 Best Local Similarity 78.6%; Pred. No. 3.6e-91;  
 Matches 305; Conservative 30; Mismatches 52; Indels 1;  
 OY 1. LHHYSLTHHEHLLLPKLNKLNRLRLRQNDNSRVPLREELDDEGSGVGGV 90  
 DB 3 LHHYSLTHHEHLLLPKLNKLNRLRLRQNDNSRVPLREELDDEGSGVGGV 62  
 OY 91 MGNKLVVHPPEKVVVDIDNDITRTTSFVALRNSVVLHVLVPKDPVLMK 150  
 DB 91 MGNKLVVHPPEKVVVDIDNDITRTTSFVALRNSVVLHVLVPKDPVLMK 150  
 OY 63 MSTRKLVVHPPEKVVVDIDNDITRTTSFVALRNSVVLHVLVPKDPVLMK 122  
 DB 63 MSTRKLVVHPPEKVVVDIDNDITRTTSFVALRNSVVLHVLVPKDPVLMK 122  
 OY 151 VEVYDSTYDMGLDQDIKEIKVEYELPTGHELFPSLGIHQKGVLLGPGPGTLL 210  
 DB 123 VEVYDSTYDMGLDQDIKEIKVEYELPTGHELFPSLGIHQKGVLLGPGPGTLL 182  
 OY 211 AAADNHTCTCFVNGSGELVQKTIQSGRWRELFPMAEHAHSTIIPMDLISGAAM 270  
 DB 183 AAADNHTCTCFVNGSGELVQKTIQSGRWRELFPMAEHAHSTIIPMDLISGAAM 242  
 OY 271 EESGSGNSRPMPTFLALQDQGEPMKLYVMARITDIOALPSEIDPKRIFP 301  
 DB 243 E-SSGSGNSRPMPTFLALQDQGEPMKLYVMARITDIOALPSEIDPKRIFP 301  
 OY 331 TPVSESLDILHHSRPMVNGDILKTAIPVNGSGELVQKTIQSGRWRELFPMAE 390  
 DB 302 TPVSESLDILHHSRPMVNGDILKTAIPVNGSGELVQKTIQSGRWRELFPMAE 361  
 OY 391 TQDFPEANQILHHSRPMVNGDILKTAIPVNGSGELVQKTIQSGRWRELFPMAE 418  
 DB 362 TQDFPEANQILHHSRPMVNGDILKTAIPVNGSGELVQKTIQSGRWRELFPMAE 389  
 RESULT 14  
 OY 097265 PRELIMINARY; PRT; 411 AA.  
 DB 01-MAY-2000 (TrEMBLrel, 13, Created)  
 DT 01-OCT-2001 (TrEMBLrel, 18, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel, 23, Last annotation update)  
 OS Hypothetical 45.7 kDa protein.  
 OC Eukaryota; Metazoa; Mammalia; Chordata; Rhodostida; Rhodostidae;  
 OC Rhodostida; Pelecaninae; Caenothobididae;  
 RN [1]  
 RM [1]

RP SEQUENCE FROM N.A.  
 RX STRAIN-Bxistol N2;  
 RX MBLINB-99069613; PubMed:9851916;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology." The C. elegans Sequencing Consortium";  
 RL Science 282:2012-2013(1998).  
 RP SEQUENCE FROM N.A.  
 RA Latreille P., Kramar J., Kappeler J.,  
 RA Submitted (Oct-1998) to the EMBL/Genbank/DBJ databases.  
 RC STRAIN-Bxistol N2;  
 RL The sequence of C. elegans strain Bxistol N2.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bxistol N2;  
 RL Submitted (Jul-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF099922; GenBank: U07212.  
 DR MBLINB: F56711.49; GenBank: F56711.49;  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR003590; AAA\_sub.  
 DR SMART: SM00383; AAA; 1.  
 DR TIGRFAMs: TIGR01242; 265945; 1.  
 DR PROSITE: PS00674; AAA; 1.  
 DR ATTBinding: protease.  
 SQ 5500000 411 AA; 45099 MW; E53P75661H7926C3 CR664;  
 Query Match 70.9%; Score 1501; DB 5; Length 411;  
 Best Local Similarity 72.5%; Pred. No. 2.3e-89;  
 Matches 292; Conservative 53; Mismatches 52; Indels 6; Gaps 2;  
 OY 21 SAMP----TQCGSLHRYVGLMHEHQLLRQTNHNLRLRQNDNSRVPLREELQ 75  
 DB 10 SAMP----TQCGSLHRYVGLMHEHQLLRQTNHNLRLRQNDNSRVPLREELQ 69  
 OY 76 LQPSGVGEVGVVWGVKLVVHPPEKVVVDIDNDITRTTSFVALRNSVVLHVLVPK 135  
 DB 76 LQPSGVGEVGVVWGVKLVVHPPEKVVVDIDNDITRTTSFVALRNSVVLHVLVPK 129  
 OY 136 LVPSPVPMVLMKVEKVPDSTYDMGLDQDIKEIKVEYELPTGHELFPSLGIHQK 185  
 DB 136 LVPSPVPMVLMKVEKVPDSTYDMGLDQDIKEIKVEYELPTGHELFPSLGIHQK 185  
 OY 190 GVLYGPGPGTLLAAVYHHTCTCFVNGSGELVQKTIQSGRWRELFPMAEHAHST 249  
 DB 190 GVLYGPGPGTLLAAVYHHTCTCFVNGSGELVQKTIQSGRWRELFPMAEHAHST 249  
 OY 256 IIPMDLISGAAMSG 315  
 DB 256 IIPMDLISGAAMSG 308  
 OY 316 ALAPRQITDKITFPNRESRLDILHHSRPMVNGDILKTAIPVNGSGELVQKTIQSG 375  
 DB 309 ALAPRQITDKITFPNRESRLDILHHSRPMVNGDILKTAIPVNGSGELVQKTIQSG 366  
 OY 376 TQDFPEANQILHHSRPMVNGDILKTAIPVNGSGELVQKTIQSGRWRELFPMAE 418  
 DB 369 TQDFPEANQILHHSRPMVNGDILKTAIPVNGSGELVQKTIQSGRWRELFPMAE 411  
 RESULT 15  
 OY 098061 PRELIMINARY; PRT; 432 AA.  
 DB 01-MAR-2002 (TrEMBLrel, 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel, 27, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel, 23, Last annotation update)  
 OS Hypothetical protein F56711.4b.

[illegible]